

An order-theoretic characterization of the Howard-Bachmann-hierarchy

Jeroen Van der Meeren*, Michael Rathjen†, Andreas Weiermann‡

Abstract

In this article we provide an intrinsic characterization of the famous Howard-Bachmann ordinal in terms of a natural well-partial-ordering by showing that this ordinal can be realized as a maximal order type of a class of generalized trees with respect to a homeomorphic embeddability relation. We use our calculations to draw some conclusions about some corresponding subsystems of second order arithmetic. All these subsystems deal with versions of light-face Π_1^1 -comprehension.

1 Introduction

The famous Howard-Bachmann ordinal η_0 (which in the literatur is also denoted by $\psi_{\varepsilon_{\Omega+1}}$, $\vartheta_{\varepsilon_{\Omega+1}}$, $\theta_{\varepsilon_{\Omega+1}0}$, $d\varepsilon_{\Omega+1}$) belongs to the most well-established arsenal of proof-theoretic ordinals of natural theories for developing significant parts of (impredicative) mathematics. Of course η_0 is much bigger than ε_0 , the proof-theoretic ordinal of first order Peano arithmetic, and it is also bigger than Γ_0 , the proof-theoretic ordinal of predicative analysis. The ordinal η_0 is the proof-theoretic ordinal of the first order theory ID_1 , which extends PA by schemes for smallest fixed points of non-iterated positive inductive definitions. The ordinal η_0 is also the proof-theoretic ordinal of the theory $KP\omega$ which formalizes an admissible universe containing ω , and η_0 is also the proof theoretic ordinal of

*jeroen.vandermeeren@ugent.be, Department of Mathematics, Ghent University, Krijgslaan 281, 9000 Gent, Belgium

†rathjen@maths.leeds.ac.uk, Department of Pure Mathematics, University of Leeds, Leeds LS2 9JT, England

‡andreas.weiermann@ugent.be, Department of Mathematics, Ghent University, Krijgslaan 281, 9000 Gent, Belgium

$ACA_0 + (\Pi_1^1 - CA)^-$ which formalizes lightface Π_1^1 -comprehension and of the theory $RCA_0 + (BI)$ which extends RCA_0 by a scheme of bar induction.

All these theories are considered to be impredicative. For example the theory $ACA_0 + (\Pi_1^1 - CA)^-$ allows the formation of new sets of natural numbers by using a comprehension formula which may contain a set quantifier ranging over the set which is just defined using the comprehension under consideration. An ordinal analysis for each of these impredicative theories turned out to be difficult. In technical terms this is usually reflected by proof calculi with rules where the complexity of the antecedent is greater than the complexity of the succedent. Typically this is a stumbling block for cut elimination and only very sophisticated methods like Buchholz' operator-controlled derivations can circumvent it.

It is important to notice that proof-theoretic ordinals are more than just set-theoretic ordinals. They usually come equipped with a first order structure with certain built-in functions which generate the ordinal in question in some natural and perspicuous way. On a more combinatorial level such an ordinal is then usually represented by a certain primitive recursive set of terms with a primitive recursive well-ordering relation. A precise description on what the characteristics of a proof-theoretic ordinal are runs under the "natural well-ordering problem" which is known to be notoriously difficult. Investigations on proof-theoretic ordinals have been undertaken by many people. To name a few: Aczel, Arai, Bachmann, Bridges, Buchholz, Feferman, Gerben, Girard, Gordeev, Isles, Jäger, Kino, Levitz, Okada, Pfeiffer, Pohlers, Probst, Rathjen, Schütte, Setzer, Strahm, Takeuti, Veblen and Weiermann.

A very important aspect of investigations on proof-theoretic ordinals goes back to Diana Schmidt who first recognized the connection between these ordinals and order-theoretic properties of the functions generating them. She characterized completely the order types which could be generated from the ordinal 0 by applying a monotonic increasing function. A monotonic increasing binary function generates out of the singleton set containing the ordinal 0 no order type larger than ε_0 . Functions of bigger arities produce easily ordinals bigger than Γ_0 and in fact ordinals of size comparable to the small Veblen ordinal $\vartheta\Omega^\omega$ but by no means an ordinal which is of size comparable to η_0 .

Diana Schmidt moreover showed that studying bounds on closure ordinals can best be achieved by determining maximal order types of well-partial-orderings which reflect monotonicity properties of the functions in question. With regard to this research program she classified maximal order types for various classes of labelled trees. The ordinals obtained in this way are all around the small Veblen ordinal $\vartheta\Omega^\omega$ and way below η_0 and it was for some time not clear whether η_0 can be characterized in terms of closure ordinals.

In Weiermann [25], it was shown that η_0 could indeed be characterized as a closure ordinal of so-called essentially monotonic increasing functions. Since then it has been open whether a corresponding order-theoretic characterization in terms of maximal order types is possible. Weiermann's proof made essential use of the linearity of ordinals and did not generalize to partially ordered structures.

Extending Schmidt's work in [26], the third author provided in a first step, an order-theoretic characterization for the large Veblen ordinal $\vartheta\Omega^\Omega$. Quite recently, the authors of this paper were able to provide in [23] much more convincing methods and results which were suitable for being extended to larger ordinals as well. This recent approach is already far reaching but still misses essential ingredients for an order-theoretic characterization of η_0 . This paper will contribute to this problem.

It should also be noticed that H. Friedman defined in 1985 tree-embeddability relations with a so-called gap-condition which generated ordinals of size $\psi\Omega_\omega$, an ordinal which is much bigger than η_0 . So in principal it seemed plausible that it is possible to single out a natural subordering of Friedman's ordering which would match with η_0 . This paper provides also a positive answer to this challenge: the defined well-partial-order of maximal order type η_0 can be seen as a natural subordering of Friedman's tree-ordering.

Moreover, we believe that our analysis will be a starting point for classifying the maximal order types of the full Friedman's gap-ordering on trees. We expect that analyzing Friedman's embeddability relations will be rather hard and difficult and we hope that the result provided in this paper yields a roadmap for a more general result. We confine ourselves to η_0 since this ordinal is somehow the first serious step into impredicativity. To elaborate a little bit more on Friedman's gap-ordering: we believe that the maximal order type of Friedman's trees on n labels can be described using the n^{th} regular cardinal number Ω_n . Hence, the uncountable cardinal numbers play a very important role in classifying the strength of Friedman's well-partial-order.

In section 2, we give some preliminaries that are needed for later sections. In section 4, we yield a well-partial-order of maximal order type η_0 , that can be seen as a natural subordering of Friedman's tree-ordering. Section 3 is needed to obtain the results in section 4. In the last two sections, section 5 and 6, results from a pure proof-theoretical point of view are studied. We determine bounds on the proof-theoretical ordinals of theories corresponding to our well-partial-orders, from which we obtain unprovability results concerning the well-partial-orderedness of these partial orders. All these theories deal with versions of light-face Π_1^1 -comprehension.

2 Preliminaries

2.1 Well-partial-orderings

Well-partial-orderings are ordered structures that are used in different fields of mathematics. For example in Gröbnerbases [1] and rewrite theory [24]. Moreover, they are well-known objects among logicians. Well-partial-orderings can be seen as generalizations of well-orderings, an important notion used in ordinal analysis [3, 15, 16]. They are the underlying concepts of the theorem of Higman [7], the theorem of Kruskal [12], Fraïssé's order type conjecture [13] and Friedman's gap-embeddability relation on trees [21]. In [17], the second and third author did a complete proof-theoretical analysis of the theorem of Kruskal.

Definition 2.1. A *well-partial-ordering* (hereafter wpo) is a partial ordering (X, \leq_X) such that for every infinite sequence $(x_i)_{i=1}^{+\infty}$ of elements in X , there exists two indices i and j such that $i < j$ and $x_i \leq_X x_j$. We denote the $\text{wpo}(X, \leq_X)$ by X if the ordering is clear from the context.

So a wpo is a well-founded partial ordering that does not admit an infinite antichain. Equivalently, it is a partial order such that every extension is well-founded. In the literature, one is more familiar with the notion of a well-quasi-order. That is a quasi-order (no antisymmetry) with the same distinctive property. However, by canceling out an obvious equivalence relation, one gets a well-partial-order. Therefore, we can restrict ourselves only to wpo 's. The interested reader can read more about well-partial-ordering in [11], the most ground-breaking paper on this subject.

Definition 2.2. The *maximal order type* of the $\text{wpo}(X, \leq_X)$ is defined as

$$\sup\{\alpha: \preceq \text{ is an extension of } \leq_X, \preceq \text{ is a well-ordering on } X \\ \text{and } \text{otype}(X, \preceq) = \alpha\}.$$

We denote this ordinal as $o(X, \leq_X)$ or as $o(X)$ if the ordering on X is clear from the context.

The maximal order type of a wpo is an important characteristic of that wpo . E.g. one can use it in determining the exact proof-theoretical strength of the wpo under consideration. In [11], it is proved that this supremum is actually a maximum, meaning that every $\text{wpo } X$ has at least one maximal linear extension.

Theorem 2.3 (de Jongh and Parikh [11]). *Assume that (X, \leq_X) is a wpo . Then there exists a well-ordering \preceq on X which is an extension of \leq_X and $\text{otype}(X, \preceq) = o(X, \leq_X)$.*

In this paper, we are interested in studying the maximal order type of specific wpo's. The technique of reifications is very useful in obtaining upper bounds [20] and lower bounds can be acquired from finding linearizations. Our technique will use the concept of left-sets $L(x)$ and quasi-embeddings. Of course, this is also interwoven with the previous mentioned techniques.

Definition 2.4. Let (X, \leq_X) be a partial order and $x \in X$. Define the left set $L_X(x)$ as the set $\{y \in X : x \not\leq_X y\}$ and $l_X(x) := o(L_X(x))$. We skip the subscript X if this is clear from the context.

Theorem 2.5 (de Jongh and Parikh [11]). Assume that X is a partial ordering. If $L(x)$ is a wpo for every $x \in X$, then X is a wpo. (The converse is trivially true.) In this case, $o(X) = \sup\{l(x) + 1 : x \in X\}$.

Using this result, it can be easily seen that the maximal order type is equal to the height of the root in the tree of finite bad sequences. If one would have that $l(x) < \alpha$, for every $x \in X$, then $o(X) \leq \alpha$. Therefore, we are really interested in characterizing the left-sets of a wpo. This, in combination with the notion of quasi-embeddings, will be the most important building blocks of the proofs in sections 3 and 4.

Definition 2.6. Let X and Y two posets. A map $e : X \rightarrow Y$ is called a **quasi-embedding** if for all $x, x' \in X$ with $e(x) \leq_Y e(x')$ we have $x \leq_X x'$.

Lemma 2.7. If X and Y are posets and $e : X \rightarrow Y$ is a quasi-embedding and Y is a wpo, then X is a wpo and $o(X) \leq o(Y)$.

2.2 Constructions on Well-partial-orderings

If we have wpo, one can construct plenty of other well-partial-orderings from it. The two most important examples are disjoint unions and products.

Definition 2.8. Let X_0 and X_1 be two partial orders. Define the **disjoint union** $X_0 + X_1$ as the set $\{(x, 0) : x \in X_0\} \cup \{(y, 1) : y \in X_1\}$ with the following ordering:

$$(x, i) \leq (y, j) \Leftrightarrow i = j \text{ and } x \leq_{X_i} y.$$

We notate the element (x, i) as x if it is clear from the context in which set x lies in. Define the **cartesian product** $X_0 \times X_1$ as the set $\{(x, y) : x \in X_0, y \in X_1\}$ with the following ordering:

$$(x, y) \leq (x', y') \Leftrightarrow x \leq_{X_0} x' \text{ and } y \leq_{X_1} y'.$$

With X^n we denote the partial ordering $X \times \cdots \times X$, where X occurs n times.

In [7], Higman studied one of the most well-known constructor on well-partial-orderings. It is in some sense a cornerstone in the theory of wpo's.

Definition 2.9. Let (X, \leq_X) be a partial order. Define (X^*, \leq_X^*) as the partial ordering on the set X^* of **finite sequences over X** ordered by

$$(x_1, \dots, x_n) \leq_X^* (y_1, \dots, y_m) \\ \Leftrightarrow (\exists 1 \leq i_1 < \dots < i_n \leq m)(\forall j \in \{1, \dots, n\})(x_j \leq_X y_{i_j}).$$

We notate this partial ordering also as (X^*, \leq^*) or even as X^* .

In the next theorem, we state the connection of the maximal order type of these constructors on well-partial-orderings X and the original order type $o(X)$.

Theorem 2.10 (de Jongh and Parikh[11], Schmidt[18]). *If X_0 , X_1 and X are wpo's, then $X_0 + X_1$, $X_0 \times X_1$ and X^* are still wpo's, and*

$$o(X_0 + X_1) = o(X_0) \oplus o(X_1), \\ o(X_0 \times X_1) = o(X_0) \otimes o(X_1),$$

where \oplus and \otimes is the natural sum and product between ordinals, and

$$o(X^*) = \begin{cases} \omega^{\omega^{o(X)}-1} & \text{if } o(X) \text{ is finite,} \\ \omega^{\omega^{o(X)}+1} & \text{if } o(X) = \varepsilon + n, \text{ with } \varepsilon \text{ an epsilon number and } n < \omega, \\ \omega^{\omega^{o(X)}} & \text{otherwise.} \end{cases}$$

In section 4, we will prove that a specific ordering $\mathcal{T}(\mathcal{B}(\cdot))$ is a well-partial-ordering with order type equal to the Howard-Bachman number. This ordering can be seen as a subordering of Friedman's famous ordering on finite trees and it uses the following constructor on wpo's.

Definition 2.11. Let X be a partial order. Define $\mathcal{B}(X)$ as the partial order where the underlying set is the set of the finite structured binary trees with leaf-labels in X and with the usual embeddability relation between trees. The internal nodes do not have labels. This means that if a tree $B \in \mathcal{B}(X)$ is embeddable in a tree $B' \in \mathcal{B}(X)$, then either B and B' are trees of one node with label $x \in X$ and $x' \in X$ respectively with $x \leq_X x'$ or B' is a binary tree with a left immediate subtree B'_1 and a right immediate subtree B'_2 and B is embeddable in B'_1 or B'_2 or B has also a left immediate subtree B_1 and a right immediate subtree B_2 and B_i is embeddable in B'_i for $i = 1, 2$. This implies that if B is embeddable in B' , then the internal nodes of B are mapped on the internal nodes of B' and the leaf-nodes of B on the leaf-nodes of B' .

One can prove that this construct a well-partial-ordering starting from a wpo X . We are interested in its maximal order type.

Definition 2.12. Define $\varphi_0\beta$ as ω^β and let φ_α be the enumeration function of the common fixed points of all φ_γ with $\gamma < \alpha$. This constructs the so-called Veblen hierarchy.

Definition 2.13. Let α be an ordinal.

$$\bar{\alpha} := \begin{cases} \alpha - 1 & \text{if } \alpha \text{ is finite,} \\ \alpha + 1 & \text{if } \alpha = \varphi_2\beta + n \text{ with } n \text{ a natural number,} \\ \alpha & \text{otherwise.} \end{cases}$$

For a proof of the following theorem, we refer the reader to [18].

Theorem 2.14. If X is a wpo, then $\mathcal{B}(X)$ is a wpo and $o(\mathcal{B}(X)) = \varepsilon_{\overline{o(X)}}$.

2.3 Generalized tree-structures

In section 4, we present a wpo of order type η_0 . This will be a tree-structure that can be interpreted as a subordering of Friedman's famous wpo on trees with gap-condition. In section 3, we give some auxiliary wpo's with maximal order types strictly below η_0 . These wpo's can also be seen as tree-representations of ordinals below η_0 . In this subsection, we give the definitions of these wpo's. The reader can also find this kind of wpo's in [23] and [27]. For the actual proofs of well-partial-orderedness and the maximal order types, the reader has to wait until sections 3 and 4. Before we present the definition, we elaborate on the definition of theta-functions.

Definition 2.15. Let Ω denote the first uncountable ordinal. Every ordinal $0 < \alpha < \varepsilon_{\Omega+1}$ can be written as $\Omega^{\alpha_1}\beta_1 + \dots + \Omega^{\alpha_n}\beta_n$ with $\beta_i < \Omega$ and $\alpha > \alpha_1 > \dots > \alpha_n$. Define the set of coefficients recursively as $K(\alpha) = \{\beta_1, \dots, \beta_n\} \cup K(\alpha_1) \cup \dots \cup K(\alpha_n)$. Let $K(0)$ be $\{0\}$. Define then $k(\alpha)$ as the ordinal $\max(K(\alpha))$.

Definition 2.16. For an ordinal α , define $\Omega_0[\alpha]$ as α and $\Omega_{n+1}[\alpha]$ as $\Omega^{\Omega_n[\alpha]}$.

Definition 2.17. Let P denote the set of the additive closed ordinal numbers $\{\omega^\alpha : \alpha \in ON\}$. For every ordinal $\alpha < \varepsilon_{\Omega+1}$, define $\vartheta(\alpha)$ as $\min\{\zeta \in P : k(\alpha) < \zeta \text{ and } \forall \beta < \alpha (k(\beta) < \zeta \rightarrow \vartheta(\beta) < \zeta)\}$. The Howard-Bachmann ordinal number is defined as $\eta_0 = \vartheta(\varepsilon_{\Omega+1}) = \sup_n (\vartheta(\Omega_n[1]))$.

For more information about the theta-function and its connection with Buchholz' Ψ -function, we refer the reader to [17]. There, they introduced the ϑ -function in a different way, but one can prove that they coincide with our definition if the argument is above Ω^2 . It can be shown by an easy cardinality argument that $\vartheta\alpha < \Omega$.

Lemma 2.18. $\vartheta\alpha < \vartheta\beta \iff \begin{cases} \alpha < \beta \text{ and } k(\alpha) < \vartheta\beta \\ \beta < \alpha \text{ and } \vartheta\alpha \leq k(\beta). \end{cases}$

We need the following two additional lemmas. The proofs are rather straightforward.

Lemma 2.19. *Suppose α and β are ordinals beneath $\varepsilon_{\Omega+1}$. Then*

$$\begin{aligned} k(\alpha \oplus \beta) &\leq k(\alpha) \oplus k(\beta), \\ k(\alpha \otimes \beta) &\leq \max\{k(\alpha) \oplus k(\beta), k(\alpha) \otimes k(\beta) \otimes \omega\}, \\ k(\omega^\alpha) &\leq \omega^{k(\alpha)}. \end{aligned}$$

Furthermore, $k(\alpha), k(\beta) \leq k(\alpha \oplus \beta)$ and $k(\alpha) \leq k(\alpha \otimes \beta)$ if $\beta > 0$.

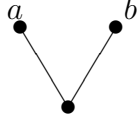
Lemma 2.20. *Suppose $\alpha_n, \dots, \alpha_0$ are countable ordinal numbers with $\alpha_i < \gamma$ for an epsilon number γ . Then $k(o((\Omega^n \alpha_n + \dots + \Omega \alpha_1 + \alpha_0)^*)) < \gamma$.*

Before we give the definition of the wpo's that we use in sections 3 and 4, let us define a specific class of constructors.

Definition 2.21. *Define Map as the least set satisfying the following:*

1. $\cdot \in \text{Map}$, (\cdot plays the role of a place holder).
2. If X is a countable wpo, then $X \in \text{Map}$,
3. If $W_1, W_2 \in \text{Map}$, then $W_1 + W_2$, $W_1 \times W_2$, W_1^* and $\mathcal{B}(W_1)$ are also elements of Map.

Every element W of Map can be seen as a mapping from the set of partial orderings to the set of partial orderings: $W(X)$ is a partial ordering by putting the partial order X into the \cdot . For example, if $W = (\mathcal{B}(\cdot) \times X)^*$, then $W(X)$ is the partial ordering $(\mathcal{B}(X) \times X)^*$. Furthermore, if X is a wpo, then $W(X)$ is a wpo and if X is countable, then so is $W(X)$. Every element of $W(X)$ is represented by a term in finitely many elements in X . For example in the case $W = (\mathcal{B}(\cdot) \times X)^*$, the element $((B(x_1, x_2), \mathfrak{x}_1), (B(x_3, x_1), \mathfrak{x}_2))$ in $W(X)$ with $x_1, x_2, x_3 \in X$, $\mathfrak{x}_1, \mathfrak{x}_2 \in X$ and $B(a, b)$ the binary tree



is represented by a term in x_1, x_2, x_3 . By deleting all entries of X , we get the naked term $((B(\cdot, \cdot), \mathbf{x}_1), (B(\cdot, \cdot), \mathbf{x}_2))$, which we notate as $w(\cdot, \cdot, \cdot)$, where

$$w(a, b, c, d) = ((B(a, b), \mathbf{x}_1), (B(c, d), \mathbf{x}_2)).$$

Therefore, the element $((B(x_1, x_2), \mathbf{x}_1), (B(x_3, x_1), \mathbf{x}_2))$ can be described using this naked term $w(\cdot, \cdot, \cdot)$ and the elements $x_1, x_2, x_3 \in X$ as $w(x_1, x_2, x_3, x_1)$. In general, an element of $W(X)$ is represented as $w(x_1, \dots, x_n)$ using a naked term $w(\cdot, \dots, \cdot)$ and elements $x_1, \dots, x_n \in X$. We will call this naked term ‘an element of W ’.

Definition 2.22. Take $W \in \text{Map}$. Define $\mathcal{T}(W)$ as the least set satisfying the following requirements

1. $\circ \in \mathcal{T}(W)$,
2. If $w(\cdot, \dots, \cdot)$ is an element of W and $t_1, \dots, t_n \in \mathcal{T}(W)$, then $w(t_1, \dots, t_n)$ is an element of $W(\mathcal{T}(W))$ and let $\circ[w(t_1, \dots, t_n)] \in \mathcal{T}(W)$. We will say that t_1, \dots, t_n have a lower complexity than $\circ[w(t_1, \dots, t_n)]$.

Let the underlying ordering $\leq_{\mathcal{T}(W)}$ be the least binary reflexive and transitive relation on $\mathcal{T}(W)$ such that

1. $\circ \leq_{\mathcal{T}(W)} t$ for every t in $\mathcal{T}(W)$,
2. if $s \leq_{\mathcal{T}(W)} t_j$ for a certain j , then $s \leq_{\mathcal{T}(W)} \circ[w(t_1, \dots, t_n)]$,
3. if $w(t_1, \dots, t_n) \leq_{W(\mathcal{T}(W), \leq_{\mathcal{T}(W)})} w'(t'_1, \dots, t'_{n'})$, then $\circ[w(t_1, \dots, t_n)] \leq_{\mathcal{T}(W)} \circ[w'(t'_1, \dots, t'_{n'})]$.

If it is clear from the context, we also notate $\leq_{\mathcal{T}(W)}$ as \leq . Sometimes, we notate the elements $\circ[w(t_1, \dots, t_n)]$ also as $\circ w(t_1, \dots, t_n)$ if $w(t_1, \dots, t_n)$ has already enough brackets in its description.

Notation 2.23. Suppose $t = \circ[w(t_1, \dots, t_n)]$ is an element of $\mathcal{T}(W)$. We denote the element $w(t_1, \dots, t_n)$ of $W(\mathcal{T}(W))$ also as $\times t$.

Our general conjecture is that for every $W \in \text{Map}$, the partial ordering $\mathcal{T}(W)$ is actually a wpo and the maximal order type is equal to $\vartheta(o(W(\Omega)))$ if $\Omega^3 \leq o(W(\Omega)) \leq \varepsilon_{\Omega+1}$. In section 3, we prove for specific $W \in \text{Map}$ that the ordering

$\mathcal{T}(W)$ is indeed a wpo and $\vartheta(o(W(\Omega)))$ is an upper bound on the maximal order type of $\mathcal{T}(W)$. In section 4, we show that $\mathcal{T}(\mathcal{B}(\cdot))$ is also a wpo and $\vartheta(\varepsilon_{\Omega+1})$ is exactly equal to $o(\mathcal{T}(\mathcal{B}(\cdot)))$. This wpo can be seen as a subordering of Friedman's famous trees with gap-embeddability relation [21]. For cases where $o(W(\Omega)) > \varepsilon_{\Omega+1}$, some little adaptations of the general formula $o(\mathcal{T}(W)) = \vartheta(o(W(\Omega)))$ are needed because the domain of the theta-function is below $\varepsilon_{\Omega+1}$. This is however beyond the scope of this article and will be treated in latter work. We believe that generalizations will lead to a full classification of the strength of Friedman's wpo's.

The next lemma is a very important lemma for the rest of the article. We will skip its proof, but one can find some subparts and the general idea of this proof in [23].

Lemma 2.24 (Lifting Lemma). *Assume that $W \in \text{Map}$ and let q be a quasi-embedding from the partial ordering Y to the partial ordering Z . Then for all elements $y_1, \dots, y_n, y'_1, \dots, y'_m$ in Y and $w(\cdot, \dots, \cdot), v(\cdot, \dots, \cdot)$ in W the inequality $w(q(y_1), \dots, q(y_n)) \leq_{W(Z)} v(q(y'_1), \dots, q(y'_m))$ implies $w(y_1, \dots, y_n) \leq_{W(Y)} v(y'_1, \dots, y'_m)$.*

Before we go further, we want to show that $\mathcal{T}(\mathcal{B}(\cdot))$ can indeed be seen as a subordering of Friedman's wpo with gap-condition. First, we give the definition of his wpo.

Definition 2.25. *Let T_n be the set of finite rooted trees with labels in $\{0, \dots, n-1\}$. An element of T_n is of the form (T, l) , where T is a finite rooted tree, which we see as a partial ordering on a set of nodes, and l is a labeling function, a mapping from T to the set $\{0, \dots, n-1\}$. Define $(T_1, l_1) \leq_{\text{gap}} (T_2, l_2)$ if there exists an injective order- and infimum-preserving mapping f from T_1 to T_2 such that*

1. $\forall \tau \in T_1$, we have $l_1(\tau) = l_2(f(\tau))$.
2. $\forall \tau \in T_1$ and for all immediate successors $\tau' \in T_1$ of τ , we have that if $\bar{\tau} \in T_2$ and $f(\tau) < \bar{\tau} < f(\tau')$, then $l_2(\bar{\tau}) \geq l_2(f(\tau')) = l_1(\tau')$.

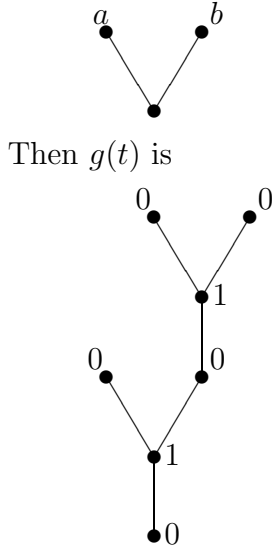
In [21], this is the so-called *weak* gap-embeddability relation.

Definition 2.26. *Define the partial ordering \bar{T}_2 as the subset of (T_2, \leq_{gap}) which consists of all finite rooted trees such that nodes with label 0 has zero or one immediate successor(s) and nodes with label 1 has exactly two immediate successors. Furthermore, every tree in \bar{T}_2 has a root with label 0.*

Note that the trees in T_n are unstructured. The trees in $\mathcal{T}(\mathcal{B}(\cdot))$ are structured, so to see the resemblance between $\mathcal{T}(\mathcal{B}(\cdot))$ and \overline{T}_2 we have to restrict \overline{T}_2 even a little bit more: we say that every tree in \overline{T}_2 is structured, meaning that it has a left-hand side and a right-hand side and an embedding between two trees preserves these left-right order.

Lemma 2.27. *The partial-ordering $\mathcal{T}(\mathcal{B}(\cdot))$ is order-isomorphic to the partial ordering \overline{T}_2 .*

Proof. Define $g : \mathcal{T}(\mathcal{B}(\cdot)) \rightarrow \overline{T}_2$ as follows. Let $g(\circ)$ be the tree which consists of one node with label 0. Take $t = \circ[B(t_1, \dots, t_n)]$ with $B(t_1, \dots, t_n)$ a binary tree with leaf-labels in the set $\{t_1, \dots, t_n\}$ and assume that $g(t_1), \dots, g(t_n)$ is already defined. Set $g(t)$ then as the tree consisting of a root with label 0, that root connected with an edge to the root of $B(t_1, \dots, t_n)$. Give all the internal nodes of B label 1 and plug $g(t_i)$ in the leaves of $B(t_1, \dots, t_n)$ with label t_i for every i . For example, if $t = \circ[B(\circ, \circ[B(\circ, \circ)])]$, with $B(a, b)$ equal to



It is easy to see that g is surjective. If we can prove that

$$t \leq_{\mathcal{T}(\mathcal{B}(\cdot))} t' \Leftrightarrow g(t) \leq_{\overline{T}_2} g(t'),$$

we are done. We will prove this by induction on the sum of complexities of t and t' . If $t = \circ$ or $t' = \circ$, then this is trivial. Assume both t and t' are different from \circ . Let $t = \circ[B(t_1, \dots, t_n)]$ and $t' = \circ[B'(t'_1, \dots, t'_m)]$. If $t \leq_{\mathcal{T}(\mathcal{B}(\cdot))} t'$, then either $t \leq_{\mathcal{T}(\mathcal{B}(\cdot))} t'_i$ for a certain i or $B(t_1, \dots, t_n) \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} B'(t'_1, \dots, t'_m)$. In both cases, the induction hypothesis yields $g(t) \leq_{\overline{T}_2} g(t')$ quite easily. Now assume $g(t) \leq_{\overline{T}_2} g(t')$. We know that the root of $g(t)$, which has label 0, is mapped on a node with

label 0. If it is not mapped onto the root of $g(t')$, then it is mapped onto a node with label 0 in $g(t'_i)$ for a certain i . Hence $g(t) \leq_{\overline{T}_2} g(t'_i)$, so $t \leq_{\mathcal{T}(\mathcal{B}(\cdot))} t'_i \leq_{\mathcal{T}(\mathcal{B}(\cdot))} t'$. Now assume that the root of $g(t)$ is mapped onto the root of $g(t')$. Every internal node a of B has to be mapped on an internal node of B' , because otherwise the label 0 of the root of the $g(t_i)$ in which the internal node a of B is mapped, gives a contradiction with the gap-condition. Furthermore, every leaf of B , in which $g(t_i)$ are plugged in, has label 0 and is mapped on a node in $g(t')$ with label 0. We can conclude that $B(g(t_1), \dots, g(t_n)) \leq_{\mathcal{B}(\overline{T}_2)} B'(g(t'_1), \dots, g(t'_m))$. The induction hypothesis yields that g is a quasi-embedding from the set $\{t_1, \dots, t_n, t'_1, \dots, t'_m\}$ to \overline{T}_2 . So the Lifting Lemma implies $B(t_1, \dots, t_n) \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} B'(t'_1, \dots, t'_m)$. Hence, $t \leq_{\mathcal{T}(\mathcal{B}(\cdot))} t'$. \square

From the previous lemma, one can actually already conclude that $\mathcal{T}(\mathcal{B}(\cdot))$ is a wpo. Therefore, one can think that the well-partial-orderedness proof of $\mathcal{T}(\mathcal{B}(\cdot))$ in Theorem 4.1 is superfluous. However, this well-partial-orderedness proof does not need an extra argument: it follows from the calculation of an upper bound of the maximal order type of $\mathcal{T}(\mathcal{B}(\cdot))$. Therefore, we do not really waste efforts by stating it in Theorem 4.1.

3 Tree-structures below the Howard-Bachmann ordinal

In section 4, we will show that $\mathcal{T}(\mathcal{B}(\cdot))$ is a wpo with maximal order type $\vartheta(\varepsilon_{\Omega+1})$. For obtaining these results, we need to approximate this wpo. This is done from ‘below’ and is treated in this section. The next theorems are generalizations of Theorems 9 and 10 in [23]. The proofs follow the same procedures as in that article, but they are more involved.

Theorem 3.1. *Suppose $Y_{i,j,k}$ and Z_i are countable wpo’s for all indices. If*

$$W(X) = \sum_{i=0}^N \left(\left(\sum_{j=0}^{k_{i,1}} Y_{i,j,1} \times X^j \right)^* \times \dots \times \left(\sum_{j=0}^{k_{i,n_i}} Y_{i,j,n_i} \times X^j \right)^* \times X^{m_i} \times Z_i \right)$$

then $\mathcal{T}(W)$ is a wpo and $o(\mathcal{T}(W)) \leq \vartheta(o(W(\Omega)))$.

Proof. We will prove the theorem by main induction on the ordinal $o(W(\Omega))$. If $W(X)$ is the empty wpo for every X , then the theorem follows easily. We can now assume without loss of generality that Z_i , $Y_{i,k_{i,j},j}$ are non-empty and $k_{i,j} > 0$ for all i and l . If $o(W(\Omega)) < \Omega$, then \cdot does not occur in W . Therefore, $n_i = m_i = 0$

for all i . Hence, $W(X)$ is equal to a $\text{wpo} \circ \sum_{i=0}^N Z_i =: Z$. So $\mathcal{T}(W) \cong Z \cup \{0\}$, where 0 is a new element smaller than every element in Z . This yields that $\mathcal{T}(W)$ is a wpo and $o(\mathcal{T}(W)) \leq o(Z) + 1 \leq \vartheta(o(Z)) = \vartheta(o(W(\Omega)))$.

Assume from now on that $o(W(\Omega)) \geq \Omega$. We will prove that $L(t)$ is a wpo and $l(t) < \vartheta(o(W(\Omega)))$ for every t in $\mathcal{T}(W)$. Then the theorem follows from Lemma 2.5. If $t = \circ$, then $L(t)$ is the empty wpo and $l(t) = 0 < \vartheta(o(W(\Omega)))$. Assume now

$$t = \circ((\overline{t_1}, \dots, \overline{t_{n_l}}), (t_1, \dots, t_{m_l}), z)$$

with $z \in Z_l$, $t_p \in \mathcal{T}(W)$ and $\overline{t_p}$ an element in $(\sum_{j=0}^{k_{l,p}} Y_{l,j,p} \times \mathcal{T}(W)^j)^*$, meaning

$$\overline{t_p} = (\overline{\overline{t_1^p}}, \dots, \overline{\overline{t_{r_p}^p}}),$$

with $\overline{\overline{t_q^p}} \in \sum_{j=0}^{k_{l,p}} Y_{l,j,p} \times \mathcal{T}(W)^j$. So

$$\overline{\overline{t_q^p}} = (y_{p,q}, (t_1^{p,q}, \dots, t_{v_{p,q}}^{p,q}))$$

with $y_{p,q} \in Y_{l,v_{p,q},p}$ and $v_{p,q} \leq k_{l,p}$. Assume $l(t_i), l(t_k^{i,j}) < \vartheta(o(W(\Omega)))$ and $L(t_i)$ and $L(t_k^{i,j})$ are wpo 's. We want to prove that $L(t)$ is a wpo and $l(t) < \vartheta(o(W(\Omega)))$.

Suppose s is an arbitrary element in $\mathcal{T}(W)$, different from \circ . Then

$$\begin{aligned} s &= \circ((\overline{s_1}, \dots, \overline{s_{n_{l'}}}), (s_1, \dots, s_{m_{l'}}), z'), \\ \overline{s_p} &= (\overline{\overline{s_1^p}}, \dots, \overline{\overline{s_{r_p'}^p}}), \\ \overline{\overline{s_q^p}} &= (y'_{p,q}, (s_1^{p,q}, \dots, s_{w_{p,q}}^{p,q})), \end{aligned} \tag{1}$$

with

$$\begin{aligned} z' &\in Z_{l'}, \\ y'_{p,q} &\in Y_{l',w_{p,q},p}. \end{aligned}$$

We see that $s \in L(t)$ iff $s_i \in L(t)$, $s_k^{i,j} \in L(t)$ and one of the following holds:

- $l \neq l'$,
- $l = l'$, $z' \in L_{Z_l}(z)$,
- $l = l'$, $z \leq_{Z_l} z'$, $(t_1, \dots, t_{m_l}) \not\leq (s_1, \dots, s_{m_{l'}})$,
- $l = l'$, $z \leq_{Z_l} z'$, $(t_1, \dots, t_{m_l}) \leq (s_1, \dots, s_{m_{l'}})$, $(\overline{t_1}, \dots, \overline{t_{n_l}}) \not\leq (\overline{s_1}, \dots, \overline{s_{n_{l'}}})$,

Now if (c.) holds, there must be a minimal index $m(s) \leq m_l$ such that

$$t_1 \leq s_1, \dots, t_{m(s)-1} \leq s_{m(s)-1}, t_{m(s)} \not\leq s_{m(s)}.$$

If (d.) is valid, there must be a minimal index $n(s) \leq n_l$ such that

$$\overline{t_1} \leq \overline{s_1}, \dots, \overline{t_{n(s)-1}} \leq \overline{s_{n(s)-1}}, \overline{t_{n(s)}} \not\leq \overline{s_{n(s)}}.$$

By similar arguments,

$$\overline{t_{n(s)}} = (\overline{\overline{t_1^{n(s)}}}, \dots, \overline{\overline{t_{r_{n(s)}}^{n(s)}}}) \not\leq (\overline{\overline{s_1^{n(s)}}}, \dots, \overline{\overline{s_{r_{n(s)}}^{n(s)}}}) = \overline{s_{n(s)}}$$

holds iff we are in one of the next cases

1. $\overline{\overline{t_1^{n(s)}}} \not\leq \overline{\overline{s_i^{n(s)}}}$ for every i ,
2. there exists an index i_1 such that $\overline{\overline{t_1^{n(s)}}} \not\leq \overline{\overline{s_{i_1}^{n(s)}}}$ for every $i < i_1$, $\overline{\overline{t_1^{n(s)}}} \leq \overline{\overline{s_{i_1}^{n(s)}}}$ and $\overline{\overline{t_2^{n(s)}}} \not\leq \overline{\overline{s_i^{n(s)}}}$ for every $i > i_1$,
- ...
- $r_{n(s)}$. there exist indices $i_1 < \dots < i_{r_{n(s)}-1}$ such that $\overline{\overline{t_1^{n(s)}}} \not\leq \overline{\overline{s_{i_1}^{n(s)}}}$ for every $i < i_1$, $\overline{\overline{t_1^{n(s)}}} \leq \overline{\overline{s_{i_1}^{n(s)}}}$ and $\overline{\overline{t_2^{n(s)}}} \not\leq \overline{\overline{s_i^{n(s)}}}$ for every $i_2 > i > i_1, \dots, \overline{\overline{t_{r_{n(s)}-1}^{n(s)}}} \leq \overline{\overline{s_{i_{r_{n(s)}-1}^{n(s)}}}$ and $\overline{\overline{t_{r_{n(s)}}^{n(s)}}} \not\leq \overline{\overline{s_i^{n(s)}}}$ for every $i > i_{r_{n(s)}-1}$.

Now,

$$\overline{\overline{t_i^{n(s)}}} = \left(y_{n(s),i}, \left(t_1^{n(s),i}, \dots, t_{v_{n(s),i}}^{n(s),i} \right) \right) \not\leq \left(y'_{n(s),j}, \left(s_1^{n(s),j}, \dots, s_{w_{n(s),j}}^{n(s),j} \right) \right) = \overline{\overline{s_j^{n(s)}}}$$

holds iff one of the following is valid

1. $v_{n(s),i} \neq w_{n(s),j}$,
2. $v_{n(s),i} = w_{n(s),j}$ and $y_{n(s),i} \not\leq_{\mathbb{Y}_{l,v_{n(s),i},n(s)}} y'_{n(s),j}$,
3. $v_{n(s),i} = w_{n(s),j}$, $y_{n(s),i} \leq_{\mathbb{Y}_{l,v_{n(s),i},n(s)}} y'_{n(s),j}$ and there exists a minimal index $p_{i,j}(s) \leq v_{n(s),i}$ such that

$$t_1^{n(s),i} \leq s_1^{n(s),j}, \dots, t_{p_{i,j}(s)-1}^{n(s),i} \leq s_{p_{i,j}(s)-1}^{n(s),j}, t_{p_{i,j}(s)}^{n(s),i} \not\leq s_{p_{i,j}(s)}^{n(s),j}.$$

We just completely characterized $L(t)$. Using this characterization, we define the following constructor $W'(X)$ in *Map*: let $W'(X)$ be $W'_1(X) + W'_2(X)$ with $W'_1(X)$ equal to

$$\begin{aligned} & \sum_{i=0, i \neq l}^N \left(\left(\sum_{j=0}^{k_{i,1}} Y_{i,j,1} \times X^j \right)^* \times \cdots \times \left(\sum_{j=0}^{k_{i,n_i}} Y_{i,j,n_i} \times X^j \right)^* \times X^{m_i} \times Z_i \right) \\ & + \left(\left(\sum_{j=0}^{k_{l,1}} Y_{l,j,1} \times X^j \right)^* \times \cdots \times \left(\sum_{j=0}^{k_{l,n_l}} Y_{l,j,n_l} \times X^j \right)^* \times X^{m_l} \times L_{Z_l}(z) \right) \\ & + \sum_{m=1}^{m_l} \left(\left(\sum_{j=0}^{k_{l,1}} Y_{l,j,1} \times X^j \right)^* \times \cdots \times \left(\sum_{j=0}^{k_{l,n_l}} Y_{l,j,n_l} \times X^j \right)^* \right. \\ & \quad \left. \times X^{m_l-1} \times L_{\mathcal{T}(W)}(t_m) \times Z_l \right) \end{aligned}$$

and $W'_2(X)$ equal to

$$\begin{aligned} & \sum_{n=1}^{n_l} \sum_{q=1}^{r_n} \left[\left(\sum_{j=0}^{k_{l,1}} Y_{l,j,1} \times X^j \right)^* \times \cdots \times \left(\sum_{j=0}^{k_{l,n-1}} Y_{l,j,n-1} \times X^j \right)^* \right. \\ & \quad \times \left(\sum_{j=0}^{k_{l,n+1}} Y_{l,j,n+1} \times X^j \right)^* \times \cdots \times \left(\sum_{j=0}^{k_{l,n_l}} Y_{l,j,n_l} \times X^j \right)^* \times X^{m_l} \times Z_l \\ & \quad \left. \times \left(\sum_{j=0}^{k_{l,n}} Y_{l,j,n} \times X^j \right)^{q-1} \times (W_2'^1(X))^* \times \cdots \times (W_2'^q(X))^* \right], \end{aligned}$$

with

$$\begin{aligned} W_2'^i(X) &:= \sum_{j=0, j \neq v_{n,i}}^{k_{l,n}} (Y_{l,j,n} \times X^j) \\ &+ \left(L_{Y_{l,v_{n,i},n}}(y_{n,i}) \times X^{v_{n,i}} \right) + \sum_{p=1}^{v_{n,i}} (Y_{l,v_{n,i},n} \times X^{v_{n,i}-1} \times L_{\mathcal{T}(W)}(t_p^{n,i})). \end{aligned}$$

The three cases separated by the sign $+$ in $W'_1(X)$ corresponds to the cases (a.), (b.) and (c.). The index m in the third line of $W'_1(X)$ matches with $m(s)$. $W'_2(X)$ corresponds to case (d.). The index n in $W'_2(X)$ matches with $n(s)$ and the index q corresponds to the cases $1 \dots r_{n(s)}$ in which we are for $\overline{t_{n(s)}} \not\leq \overline{s_{n(s)}}$. $W_2'^i(X)$ matches with $\overline{t_i^{n(s)}} \not\leq \overline{s_j^{n(s)}}$, where p in $W_2'^i(X)$ corresponds to $p_{i,j}(s)$.

Now, following the here-described characterization of $L(t)$ thoroughly step-by-step, one can see that there exists a mapping f from $\{w(s_1, \dots, s_n) \in W(\mathcal{T}(W)) : \circ[w(s_1, \dots, s_n)] \in L(t)\}$ into $W'(\mathcal{T}(W))$ such that if we have $f(w(s_1, \dots, s_n)) = w'(s'_1, \dots, s'_m)$, then $\{s'_1, \dots, s'_m\} \subseteq \{s_1, \dots, s_n\} \subseteq L(t)$ and if the inequality $f(w(s_1, \dots, s_n)) \leq_{W'(\mathcal{T}(W))} f(\overline{w}(\overline{s}_1, \dots, \overline{s}_k))$ holds, then $w(s_1, \dots, s_n) \leq_{W(\mathcal{T}(W))} \overline{w}(\overline{s}_1, \dots, \overline{s}_k)$. We do not explicitly write out the full details of this argument because in Theorem 4.1 we will do a similar proof (written out in full details) and that proof is less messy.

We pinpoint a mapping g from $L(t)$ into $\mathcal{T}(W')$. This mapping will be a quasi-embedding. We do this by induction on the complexity of the terms in $L(t)$. Let $g(\circ)$ be \circ . Let s be an element of $L(t)$, defined as in (1). By induction, we can assume that $g(s_i)$ and $g(s_k^{i,j})$ are already defined. We know that s is equal to

$$\circ \left[w(s_1, \dots, s_{m_{l'}}, s_1^{1,1}, \dots, s_{w_{n_{l'}, r'_{n_{l'}}}}^{n_{l'}, r'_{n_{l'}}}) \right]$$

for a certain element $w(\cdot, \dots, \cdot)$ in W . Denote

$$f(\times s) = f \left(w(s_1, \dots, s_{m_{l'}}, s_1^{1,1}, \dots, s_{w_{n_{l'}, r'_{n_{l'}}}}^{n_{l'}, r'_{n_{l'}}}) \right)$$

as $w'(s'_1, \dots, s'_m) \in W'(L(t))$ with $\{s'_1, \dots, s'_m\} \subseteq \{s_1, \dots, s_{m_{l'}}, s_1^{1,1}, \dots, s_{w_{n_{l'}, r'_{n_{l'}}}}^{n_{l'}, r'_{n_{l'}}}\}$. Define $g(s)$ as $\circ[w'(g(s'_1), \dots, g(s'_m))]$. We want to prove that the mapping g is a quasi-embedding.

We prove by induction on the sum of the complexities of s and \overline{s} that the inequality $g(s) \leq_{\mathcal{T}(W')} g(\overline{s})$ implies $s \leq_{\mathcal{T}(W)} \overline{s}$. If s or \overline{s} is equal to \circ , then this is trivial. Now, let $s = \circ[w(s_1, \dots, s_n)]$ and $\overline{s} = \circ[\overline{w}(\overline{s}_1, \dots, \overline{s}_n)]$. Assume that $f(w(s_1, \dots, s_n))$ is equal to $w'(s'_1, \dots, s'_m)$ and $f(\overline{w}(\overline{s}_1, \dots, \overline{s}_n))$ is $\overline{w}'(\overline{s}'_1, \dots, \overline{s}'_m)$. So we have

$$\circ[w'(g(s'_1), \dots, g(s'_m))] = g(s) \leq_{\mathcal{T}(W')} g(\overline{s}) = \circ[\overline{w}'(g(\overline{s}'_1), \dots, g(\overline{s}'_m))].$$

Then either $g(s) \leq_{\mathcal{T}(W')} g(\overline{s}'_q)$ for a certain q or

$$w'(g(s'_1), \dots, g(s'_m)) \leq_{W'(\mathcal{T}(W'))} \overline{w}'(g(\overline{s}'_1), \dots, g(\overline{s}'_m)).$$

In the first case, $s \leq \overline{s}'_q \leq \overline{s}$. In the latter case, the induction hypothesis yields that g is a quasi-embedding from the set $\{s'_1, \dots, s'_m, \overline{s}'_1, \dots, \overline{s}'_m\}$ to $\mathcal{T}(W')$. Hence, the Lifting Lemma implies $w'(s_1, \dots, s'_m) \leq_{W'(\mathcal{T}(W))} \overline{w}'(\overline{s}'_1, \dots, \overline{s}'_m)$. So $f(\times s) \leq f(\times \overline{s})$, hence $\times s \leq \times \overline{s}$. From this we can conclude that $s \leq \overline{s}$.

So g is a quasi-embedding from $L(t)$ in $\mathcal{T}(W')$. If $o(W'(\Omega)) < o(W(\Omega))$, the main induction hypothesis yields $\mathcal{T}(W')$ is a wpo and $o(\mathcal{T}(W')) \leq \vartheta(o(W'(\Omega)))$. Therefore, using Lemma 2.7, $L(t)$ is a wpo and $o(L(t)) \leq \vartheta(o(W'(\Omega)))$. If additionally $k(o(W'(\Omega))) < \vartheta(o(W(\Omega)))$, we can conclude that the inequality $o(L(t)) < \vartheta(o(W(\Omega)))$ holds, the objective that we want to achieve. Thus if we can prove that $o(W'(\Omega)) < o(W(\Omega))$ and $k(o(W'(\Omega))) < \vartheta(o(W(\Omega)))$, we can end the proof of this theorem.

1) $o(W'(\Omega)) < o(W(\Omega))$.

For notational convenience, we write Υ instead of $o(\Upsilon)$ for wpo 's Υ . Additionally, we write α^* instead of $o(\alpha^*)$ for ordinal numbers α .

$$\begin{aligned} o(W_2^{n_i}(\Omega)) &< \bigoplus_{j=0, j \neq v_{n,i}}^{k_{l,n}} (\Upsilon_{l,j,n} \otimes \Omega^j) \oplus (l_{\Upsilon_{l,v_{n,i},n}}(y_{n,i}) \otimes \Omega^{v_{n,i}}) \oplus \Omega^{v_{n,i}} \\ &\leq \bigoplus_{j=0}^{k_{l,n}} (\Upsilon_{l,j,n} \otimes \Omega^j). \end{aligned} \quad (2)$$

We know that $\left(\bigoplus_{j=0}^{k_{l,n}} \Upsilon_{l,j,n} \otimes \Omega^j\right)^*$ is a multiplicative closed ordinal. Therefore, inequality (2) yields

$$\left(\bigoplus_{j=0}^{k_{l,n}} \Upsilon_{l,j,n} \otimes \Omega^j\right)^{q-1} \otimes o(W_2^{n_1}(\Omega)^*) \otimes \cdots \otimes o(W_2^{n_q}(\Omega)^*) < \left(\bigoplus_{j=0}^{k_{l,n}} \Upsilon_{l,j,n} \otimes \Omega^j\right)^*.$$

From the assumption that $k_{l,n} > 0$ and $\Upsilon_{l,k_{l,n},n} \neq \emptyset$, we also have

$$\begin{aligned} Z_l \otimes \left(\bigoplus_{j=0}^{k_{l,n}} \Upsilon_{l,j,n} \otimes \Omega^j\right)^{q-1} &\otimes o(W_2^{n_1}(\Omega)^*) \otimes \cdots \otimes o(W_2^{n_q}(\Omega)^*) \\ &< \left(\bigoplus_{j=0}^{k_{l,n}} \Upsilon_{l,j,n} \otimes \Omega^j\right)^*. \end{aligned}$$

Because

$$\left(\bigoplus_{j=0}^{k_{l,1}} \Upsilon_{l,j,1} \otimes \Omega^j\right)^* \otimes \cdots \otimes \left(\bigoplus_{j=0}^{k_{l,n_l}} \Upsilon_{l,j,n_l} \otimes \Omega^j\right)^* \otimes \Omega^{m_l}$$

is additive closed, we obtain

$$o(W_2'(X)) < \left(\bigoplus_{j=0}^{k_{l,1}} \Upsilon_{l,j,1} \otimes \Omega^j\right)^* \otimes \cdots \otimes \left(\bigoplus_{j=0}^{k_{l,n_l}} \Upsilon_{l,j,n_l} \otimes \Omega^j\right)^* \otimes \Omega^{m_l}$$

and

$$\begin{aligned}
& \bigoplus_{m=1}^{m_l} \left(\left(\bigoplus_{j=0}^{k_{l,1}} Y_{l,j,1} \otimes \Omega^j \right)^* \otimes \cdots \otimes \left(\bigoplus_{j=0}^{k_{l,n_l}} Y_{l,j,n_l} \otimes \Omega^j \right)^* \right. \\
& \quad \left. \otimes \Omega^{m_l-1} \otimes L_{\mathcal{T}(W)}(t_m) \otimes Z_l \right) \oplus o(W'_2(X)) \\
& < \left(\bigoplus_{j=0}^{k_{l,1}} Y_{l,j,1} \otimes \Omega^j \right)^* \otimes \cdots \otimes \left(\bigoplus_{j=0}^{k_{l,n_l}} Y_{l,j,n_l} \otimes \Omega^j \right)^* \otimes \Omega^{m_l}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
& o(W'(\Omega)) \\
& < \bigoplus_{i=0, i \neq l}^N \left(\left(\bigoplus_{j=0}^{k_{i,1}} Y_{i,j,1} \otimes \Omega^j \right)^* \otimes \cdots \otimes \left(\bigoplus_{j=0}^{k_{i,n_i}} Y_{i,j,n_i} \otimes \Omega^j \right)^* \otimes \Omega^{m_i} \otimes Z_i \right) \\
& \oplus \left(\left(\bigoplus_{j=0}^{k_{l,1}} Y_{l,j,1} \otimes \Omega^j \right)^* \otimes \cdots \otimes \left(\bigoplus_{j=0}^{k_{l,n_l}} Y_{l,j,n_l} \otimes \Omega^j \right)^* \otimes \Omega^{m_l} \otimes l_{Z_l}(z) \right) \\
& \oplus \left(\left(\bigoplus_{j=0}^{k_{l,1}} Y_{l,j,1} \otimes \Omega^j \right)^* \otimes \cdots \otimes \left(\bigoplus_{j=0}^{k_{l,n_l}} Y_{l,j,n_l} \otimes \Omega^j \right)^* \otimes \Omega^{m_l} \right) \\
& \leq o(W(\Omega)).
\end{aligned}$$

2) $k(o(W'(\Omega))) < \vartheta(o(W(\Omega)))$.

We know that $o(W(\Omega)) \geq \Omega$, hence $\vartheta(o(W(\Omega)))$ is an epsilon number. So from Lemmas 2.19 and 2.20 the claim follows if $Y_{i,j,k}$, Z_i , $l_{\mathcal{T}(W)}(t_i)$ and $l_{\mathcal{T}(W)}(t_k^{i,j})$ are all smaller than $\vartheta(o(W(\Omega)))$. We know that this is true for $l_{\mathcal{T}(W)}(t_i)$ and $l_{\mathcal{T}(W)}(t_k^{i,j})$, by the sub-induction hypothesis. Furthermore, we have $Y_{i,j,k}, Z_k \leq k(o(W(\Omega))) < \vartheta(o(W(\Omega)))$ using Lemma 2.19. □

Theorem 3.2. *If $W(X) = X^{**}$, then $\mathcal{T}(W)$ is a wpo and $o(\mathcal{T}(W)) \leq \vartheta(o(W(\Omega))) = \vartheta(\Omega^{\Omega^{\Omega^\omega}})$.*

Proof. We show that $L(t)$ is a wpo and $l(t) < \vartheta(\Omega^{\Omega^{\Omega^\omega}})$ for every t in $\mathcal{T}(W)$ by induction on the complexity of t . The theorem then follows from Theorem 2.5. If $t = \circ$, then left-set $L(t)$ is the empty wpo and $l(t) = 0 < \vartheta(\Omega^{\Omega^{\Omega^\omega}})$. Assume

now that $t = \circ((t_1^1, \dots, t_{n_1}^1), \dots, (t_1^k, \dots, t_{n_k}^k))$. From the induction hypothesis, we know that $L(t_j^i)$ are wpo's and $l(t_j^i) < \vartheta(o(W(\Omega)))$. Assume

$$s = \circ((s_1^1, \dots, s_{m_1}^1), \dots, (s_1^l, \dots, s_{m_l}^l)).$$

Then $t \leq s$ iff $t \leq s_j^i$ for certain i and j or

$$((t_1^1, \dots, t_{n_1}^1), \dots, (t_1^k, \dots, t_{n_k}^k)) \leq^{**} ((s_1^1, \dots, s_{m_1}^1), \dots, (s_1^l, \dots, s_{m_l}^l)).$$

Hence, $s \in L(t)$ iff $s_j^i \in L(t)$ for every i and j and one of the following holds

1. $(t_1^1, \dots, t_{n_1}^1) \not\leq^* (s_1^i, \dots, s_{n_i}^i)$ for every i ,
2. there exists an index l_1 such that $(t_1^1, \dots, t_{n_1}^1) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $i < l_1$, $(t_1^1, \dots, t_{n_1}^1) \leq^* (s_1^{l_1}, \dots, s_{m_{l_1}}^{l_1})$ and $(t_1^2, \dots, t_{n_2}^2) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $i > l_1$,
3. there exist indices $l_1 < l_2$ such that $(t_1^1, \dots, t_{n_1}^1) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $i < l_1$, $(t_1^1, \dots, t_{n_1}^1) \leq^* (s_1^{l_1}, \dots, s_{m_{l_1}}^{l_1})$, $(t_1^2, \dots, t_{n_2}^2) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $l_1 < i < l_2$, $(t_1^2, \dots, t_{n_2}^2) \leq^* (s_1^{l_2}, \dots, s_{m_{l_2}}^{l_2})$ and $(t_1^3, \dots, t_{n_3}^3) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $i > l_2$,
- ...
- k. there exist indices $l_1 < \dots < l_{k-1}$ such that $(t_1^1, \dots, t_{n_1}^1) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $i < l_1$, $(t_1^1, \dots, t_{n_1}^1) \leq^* (s_1^{l_1}, \dots, s_{m_{l_1}}^{l_1})$, $(t_1^2, \dots, t_{n_2}^2) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $l_1 < i < l_2$, $(t_1^2, \dots, t_{n_2}^2) \leq^* (s_1^{l_2}, \dots, s_{m_{l_2}}^{l_2})$, ..., $(t_1^{k-1}, \dots, t_{n_{k-1}}^{k-1}) \leq^* (s_1^{l_{k-1}}, \dots, s_{m_{l_{k-1}}}^{l_{k-1}})$ and $(t_1^k, \dots, t_{n_k}^k) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $i > l_{k-1}$.

Now, $(t_1^i, \dots, t_{n_i}^i) \not\leq^* (s_1^j, \dots, s_{m_j}^j)$ is valid iff one of the following holds

1. $s_r^j \in L(t_1^i)$ for every r ,
2. there exists an index $r_1^{i,j}$ such that $s_r^j \in L(t_1^i)$ for every $r < r_1^{i,j}$, $t_1^i \leq s_{r_1^{i,j}}^j$ and $s_r^j \in L(t_2^i)$ for every $r > r_1^{i,j}$,
- ...
- n_i . there exist indices $r_1^{i,j} < \dots < r_{n_i-1}^{i,j}$ such that $s_r^j \in L(t_1^i)$ for every $r < r_1^{i,j}$, $t_1^i \leq s_{r_1^{i,j}}^j$, $s_r^j \in L(t_2^i)$ for every $r_2^{i,j} > r > r_1^{i,j}$, ..., $t_{n_i-1}^i \leq s_{r_{n_i-1}^{i,j}}^j$ and $s_r^j \in L(t_{n_i}^i)$ for every $r > r_{n_i-1}^{i,j}$.

Define

$$W'(X) := \bigoplus_{i=1}^k \left(\left(\bigoplus_{j=1}^{n_1} L(t_1^1)^* \otimes X \otimes \cdots \otimes X \otimes L(t_j^1)^* \right)^* \otimes X^* \otimes \cdots \right. \\ \left. \otimes X^* \otimes \left(\bigoplus_{j=1}^{n_i} L(t_1^i)^* \otimes X \otimes \cdots \otimes X \otimes L(t_j^i)^* \right)^* \right).$$

If $s = \circ((s_1^1, \dots, s_{m_1}^1), \dots, (s_1^l, \dots, s_{m_l}^l))$ is an element in $L(t)$, we can interpret $((s_1^1, \dots, s_{m_1}^1), \dots, (s_1^l, \dots, s_{m_l}^l))$ as an element of $W'(L(t))$. Denote this interpretation as $f(\times s) = w(s_1, \dots, s_n)$, with $\{s_1, \dots, s_n\} \subseteq \{s_1^1, \dots, s_{m_l}^l\} \subseteq L(t)$.

For example assume that case 2. holds, meaning that there exists an index l_1 such that $(t_1^1, \dots, t_{n_1}^1) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $i < l_1$, $(t_1^1, \dots, t_{n_1}^1) \leq^* (s_1^{l_1}, \dots, s_{m_{l_1}}^{l_1})$ and $(t_1^2, \dots, t_{n_2}^2) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ for every $i > l_1$. Then assume that for $(t_1^1, \dots, t_{n_1}^1) \not\leq^* (s_1^1, \dots, s_{m_1}^1)$ subcase 1. holds, but for $(t_1^1, \dots, t_{n_1}^1) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$ with $1 < i < l_1$ subcase 2. holds. Additionally, assume that for every $i > l_1$ subcase 2. holds if we look to $(t_1^2, \dots, t_{n_2}^2) \not\leq^* (s_1^i, \dots, s_{m_i}^i)$. In this case, $f(\times s) = w(s_1, \dots, s_n)$ is equal to

$$\left(\left((s_1^1, \dots, s_{m_1}^1), \left((s_1^2, \dots, s_{r_1^{1,2}-1}^2), s_{r_1^{1,2}}^2, (s_{r_1^{1,2}+1}^2, \dots, s_{m_2}^2) \right), \dots, \right. \right. \\ \left. \left((s_1^{l_1-1}, \dots, s_{r_1^{1,l_1-1}-1}^{l_1-1}), s_{r_1^{1,l_1-1}}^{l_1-1}, (s_{r_1^{1,l_1-1}+1}^{l_1-1}, \dots, s_{m_{l_1-1}}^{l_1-1}) \right) \right), \\ (s_1^{l_1}, \dots, s_{m_{l_1}}^{l_1}), \\ \left(\left((s_1^{l_1+1}, \dots, s_{r_1^{2,l_1+1}-1}^{l_1+1}), s_{r_1^{2,l_1+1}}^{l_1+1}, (s_{r_1^{2,l_1+1}+1}^{l_1+1}, \dots, s_{m_{l_1+1}}^{l_1+1}) \right), \dots, \right. \\ \left. \left((s_1^l, \dots, s_{r_1^{2,l}-1}^l), s_{r_1^{2,l}}^l, (s_{r_1^{2,l}+1}^l, \dots, s_{m_l}^l) \right) \right) \right),$$

where $w(\cdot, \dots, \cdot)$ is equal to

$$\left(\left((s_1^1, \dots, s_{m_1}^1), \left((s_1^2, \dots, s_{r_1^{1,2}-1}^2), \cdot, (s_{r_1^{1,2}+1}^2, \dots, s_{m_2}^2) \right), \dots, \right. \right. \\ \left. \left((s_1^{l_1-1}, \dots, s_{r_1^{1,l_1-1}-1}^{l_1-1}), \cdot, (s_{r_1^{1,l_1-1}+1}^{l_1-1}, \dots, s_{m_{l_1-1}}^{l_1-1}) \right) \right), \\ (\cdot, \dots, \cdot), \\ \left(\left((s_1^{l_1+1}, \dots, s_{r_1^{2,l_1+1}-1}^{l_1+1}), \cdot, (s_{r_1^{2,l_1+1}+1}^{l_1+1}, \dots, s_{m_{l_1+1}}^{l_1+1}) \right), \dots, \right. \\ \left. \left((s_1^l, \dots, s_{r_1^{2,l}-1}^l), \cdot, (s_{r_1^{2,l}+1}^l, \dots, s_{m_l}^l) \right) \right) \right),$$

Using the just-described characterization of the set $L(t)$, one can see that the inequality $f(\times s) \leq_{W'(\mathcal{T}(W))} f(\times s')$ yields $\times s \leq_{W(\mathcal{T}(W))} \times s'$ for every $s, s' \in$

$L(t) \setminus \{\circ\}$. A similar, but more detailed, argument can be found in the proof of Theorem 4.1.

Define now a mapping g from $L(t)$ into $\mathcal{T}(W')$ in the following recursive way. Let $g(\circ)$ be \circ . Assume now that $s = \circ((s_1^1, \dots, s_{m_1}^1), \dots, (s_1^l, \dots, s_{m_l}^l)) \in L(t)$ and that $g(s_j^i)$ is already defined for every i and j . If $f(\times s) = w(s_1, \dots, s_n)$, let $g(s)$ be the element $\circ[w(g(s_1), \dots, g(s_n))]$ in $\mathcal{T}(W')$. If s is the example as before, then $g(s)$ is defined as

$$\begin{aligned} & \circ \left(\left((s_1^1, \dots, s_{m_1}^1), \left((s_1^2, \dots, s_{r_1^{1,2}-1}^2), g(s_{r_1^{1,2}}^2), (s_{r_1^{1,2}+1}^2, \dots, s_{m_2}^2) \right), \dots, \right. \right. \\ & \quad \left. \left((s_1^{l_1-1}, \dots, s_{r_1^{1,l_1-1}-1}^{l_1-1}), g(s_{r_1^{1,l_1-1}}^{l_1-1}), (s_{r_1^{1,l_1-1}+1}^{l_1-1}, \dots, s_{m_{l_1-1}}^{l_1-1}) \right) \right), \\ & \quad \left(g(s_1^{l_1}), \dots, g(s_{m_{l_1}}^{l_1}) \right), \\ & \quad \left(\left((s_1^{l_1+1}, \dots, s_{r_1^{2,l_1+1}-1}^{l_1+1}), g(s_{r_1^{2,l_1+1}}^{l_1+1}), (s_{r_1^{2,l_1+1}+1}^{l_1+1}, \dots, s_{m_{l_1+1}}^{l_1+1}) \right), \dots, \right. \\ & \quad \left. \left. \left((s_1^l, \dots, s_{r_1^{2,l}-1}^l), g(s_{r_1^{2,l}}^l), (s_{r_1^{2,l}+1}^l, \dots, s_{m_l}^l) \right) \right) \right). \end{aligned}$$

Is g a quasi-embedding? We claim that $g(s) \leq_{\mathcal{T}(W')} g(s')$ implies $s \leq_{\mathcal{T}(W)} s'$ by induction on the sum of the complexities of s and s' . If $s' = \circ$ or $s = \circ$, then this is trivial. Let

$$\begin{aligned} s &= \circ((s_1^1, \dots, s_{m_1}^1), \dots, (s_1^l, \dots, s_{m_l}^l)), \\ s' &= \circ((s_1^{l_1}, \dots, s_{p_1}^{l_1}), \dots, (s_1^{l_r}, \dots, s_{p_r}^{l_r})) \end{aligned}$$

and assume that $g(s) \leq g(s')$, $f(\times s) = w(s_1, \dots, s_n)$ and $f(\times s') = w'(s'_1, \dots, s'_{n'})$. Then either $g(s) \leq g(s'_i)$ for a certain i or $\times g(s) \leq_{W'(\mathcal{T}(W'))} \times g(s')$. In the former case, the induction hypothesis yields $s \leq s'_i \leq s'$. In the latter case, we have $w(g(s_1), \dots, g(s_n)) \leq_{W'(\mathcal{T}(W))} w'(g(s'_1), \dots, g(s'_{n'}))$. The induction hypothesis yields that g is a quasi-embedding from the set $\{s_1, \dots, s_n, s'_1, \dots, s'_{n'}\}$ to $\mathcal{T}(W')$. Hence, the Lifting Lemma implies $f(\times s) = w(s_1, \dots, s_n) \leq_{W'(\mathcal{T}(W))} w'(s'_1, \dots, s'_{n'}) = f(\times s')$. From the properties of f , we obtain $\times s \leq_{W(\mathcal{T}(W))} \times s'$, hence $s \leq_{\mathcal{T}(W)} s'$.

Now, Lemma 2.7 and Theorem 3.1 yield that $L(t)$ is a wpo and

$$o(L(t)) \leq o(\mathcal{T}(W')) \leq \vartheta(o(W'(\Omega))).$$

So we can end this proof if we can show that $\vartheta(o(W'(\Omega))) < \vartheta(\Omega^{\Omega^{\Omega^\omega}})$. For notational convenience, we write Y instead of $o(Y)$ and Y^* instead of $o(Y^*)$ for

wpō's \mathbb{Y} . From the induction hypothesis, we know that $l(t_j^i) < \vartheta(\Omega^{\Omega^{\Omega^\omega}}) < \Omega$, hence $l(t_1^i)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t_j^i)^* < \Omega^n$ for a certain finite n . Therefore,

$$\left(\bigoplus_{j=1}^{n_i} l(t_1^i)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t_j^i)^* \right) + 1 < \Omega^\omega$$

and

$$\left(\bigoplus_{j=1}^{n_i} l(t_1^i)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t_j^i)^* \right)^* < \Omega^{\Omega^{\Omega^\omega}}.$$

From this it follows that $o(W'(\Omega)) < \Omega^{\Omega^{\Omega^\omega}}$. Now

$$\begin{aligned} o(W'(\Omega)) &= \bigoplus_{i=1}^k \left(\left(\bigoplus_{j=1}^{n_1} l(t_1^1)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t_j^1)^* \right)^* \otimes \Omega^* \otimes \cdots \right. \\ &\quad \left. \otimes \Omega^* \otimes \left(\bigoplus_{j=1}^{n_i} l(t_1^i)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t_j^i)^* \right)^* \right) \\ &= \bigoplus_{i=1}^k \left(\Omega^{\omega(i-1)} \otimes \left(\bigoplus_{j=1}^{n_1} \Omega^{j-1} \otimes l(t_1^1)^* \otimes \cdots \otimes l(t_j^1)^* \right)^* \otimes \cdots \right. \\ &\quad \left. \otimes \left(\bigoplus_{j=1}^{n_i} \Omega^{j-1} \otimes l(t_1^i)^* \otimes \cdots \otimes l(t_j^i)^* \right)^* \right). \end{aligned}$$

Using Lemmas 2.19 and 2.20, $k(o(W'(\Omega))) < \vartheta(\Omega^{\Omega^{\Omega^\omega}})$ is valid if

$$l(t_1^i)^* \otimes \cdots \otimes l(t_j^i)^* < \vartheta(\Omega^{\Omega^{\Omega^\omega}}).$$

This is true because $l(t_j^i) < \vartheta(\Omega^{\Omega^{\Omega^\omega}})$ and $\vartheta(\Omega^{\Omega^{\Omega^\omega}})$ is an epsilon number. \square

These results can be generalized to the following theorem. We do not give the proof as it is very technical and it will not provide new insights because it follows the same procedures as in Theorems 3.1 and 3.2.

Theorem 3.3. *Suppose that $W \in \text{Map}$ consists only of $\cdot, +, \times, *$ and countable wpō's \mathbb{X}_i . Then $\mathcal{T}(W)$ is a wpō and $o(\mathcal{T}(W)) \leq \vartheta(o(W(\Omega)))$.*

One can also show that $\vartheta(o(W(\Omega)))$ is a lower bound for $o(\mathcal{T}(W))$ if $o(W(\Omega)) \geq \Omega^3$. However, the purpose of this section is to proof theorems that are needed in section 4. Therefore, we skip the proof of the lower bound.

4 An order-theoretic approach of the Howard-Bachmann ordinal

The previous section yields

$$o(\mathcal{T}(\overbrace{*\cdots*}^n)) = \vartheta(\Omega_{2n-1}[\omega]).$$

Therefore, the tree-structures $\mathcal{T}(\overbrace{*\cdots*}^n)$ give rise to representations of countable ordinals strictly below the Howard-Bachmann ordinal and the ‘limit’ of these structures will be equal to this famous ordinal. But what do we mean by the ‘limit’ of these structures? In some sense, the set of binary trees is the limit of an iteration of the $*$ -operator. Hence, one can expect that $o(\mathcal{T}(\mathcal{B}(\cdot))) = \sup_{n < \omega} \vartheta(\Omega_{2n-1}[\omega]) = \vartheta(\varepsilon_{\Omega+1})$. In this section, we will prove that this is indeed the case. This result yields that the Howard-Bachmann ordinal can be represented as a tree-structure using binary trees, or more specifically, as the $\text{wpo}(\overline{\mathbb{T}}_2, \leq_{\text{gap}})$.

Theorem 4.1. $\mathcal{T}(\mathcal{B}(\cdot))$ is a wpo and $o(\mathcal{T}(\mathcal{B}(\cdot))) \leq \vartheta(\varepsilon_{\Omega+1})$.

Proof. We prove that $L(t)$ is a wpo and $l(t) < \vartheta(\varepsilon_{\Omega+1})$ for every t in $\mathcal{T}(\mathcal{B}(\cdot))$ by induction on the complexity of t . The theorem then follows from Theorem 2.5. If $t = \circ$, then $L(t)$ is the empty wpo and $l(t) = 0 < \vartheta(\varepsilon_{\Omega+1})$.

Let $B(t_1, \dots, t_n)$ be an element of $\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))$. If we write $B(t_1, \dots, t_n)$, we mean that the leaf-labels are elements of $\{t_1, \dots, t_n\}$. If it is clear from the context, we sometimes write B instead of $B(t_1, \dots, t_n)$. If $B(t_1, \dots, t_n)$ is a tree of height zero with leaf-label t_i , define $W_B(X)$ as the partial ordering $\mathcal{B}(L_{\mathcal{T}(\mathcal{B}(\cdot))}(t_i))$. Remark that \cdot does not occur in W_B . If $B(t_1, \dots, t_n)$ is a tree with immediate subtrees B_1 and B_2 , define $W_B(X) = W_{B(t_1, \dots, t_n)}(X)$ as

$$(W_{B_1}(X) + W_{B_2}(X))^* \times X.$$

We prove by induction on the height of the tree B that there exists a mapping g_B from

$$L_B := \{D(d_1, \dots, d_k) \in \mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot))) : B(t_1, \dots, t_n) \not\leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} D(d_1, \dots, d_k)\}$$

to the partial ordering $W_B(\mathcal{T}(\mathcal{B}(\cdot)))$ such that g_B is a quasi-embedding and if $g_B(D) = w(d'_1, \dots, d'_m)$, with w an element in W_B and $d'_1, \dots, d'_m \in \mathcal{T}(\mathcal{B}(\cdot))$, then $\{d'_1, \dots, d'_m\} \subseteq \{d_1, \dots, d_k\}$.

i) height(B) = 0.

Let $B(t_1, \dots, t_n)$ be a tree with one node and leaf-label t_i . Then $D(d_1, \dots, d_k) \in$

L_B iff $d_j \in L(t_i)$ for every j . Define then the element $g_B(D(d_1, \dots, d_k))$ as $D(d_1, \dots, d_k) \in \mathcal{B}(L_{\mathcal{T}(\mathcal{B}(\cdot))}(t_i)) = W_B(\mathcal{T}(\mathcal{B}(\cdot)))$. If we write $g_B(D)$ as $w(d'_1, \dots, d'_m)$, then $m = 0$, hence one can show easily that the desired properties of g_B are valid.

ii) $\text{height}(B) > 0$.

Let $B(t_1, \dots, t_n)$ be a binary tree with immediate subtrees B_1 and B_2 . By the induction hypothesis, there exist functions g_{B_1} and g_{B_2} with the wanted properties. Now, Pick an arbitrary $D(d_1, \dots, d_k) \in \mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))$. Then $B(t_1, \dots, t_n) \not\leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} D(d_1, \dots, d_k)$ is valid iff one of the following holds

1. $D(d_1, \dots, d_k)$ is a binary tree of height 0 with label d_i ,
2. $D(d_1, \dots, d_k)$ is a tree of height strictly larger than 0 with immediate subtrees D_1 and D_2 , $B(t_1, \dots, t_n) \not\leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} D_i$ for $i = 1, 2$ and one of the following occurs
 - (a) $B_1 \not\leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} D_1$,
 - (b) $B_1 \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} D_1$ and $B_2 \not\leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} D_2$.

Because $B(t_1, \dots, t_n) \not\leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} D_i$ for $i = 1, 2$, we can also use the above case-study for the trees D_1 and D_2 . This leads us to the following definition. Choose an arbitrary $D(d_1, \dots, d_k) \in L_B$. Define E_0 and F_0 as $D(d_1, \dots, d_k)$. Assume that we have E_i and F_i for a certain i as elements of $\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))$. If F_i is a tree of height strictly larger than 0, we want to define E_{i+1} and F_{i+1} . Suppose that F_i^1 and F_i^2 are the immediate subtrees of F_i . Then define E_{i+1} as

$$\begin{cases} F_i^1 & \text{in case that 2.(a) holds if we look to the condition } B \not\leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} F_i \\ F_i^2 & \text{in case that 2.(b) holds if we look to the condition } B \not\leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} F_i. \end{cases} \quad (3)$$

Let x_{i+1} be the number j such that $E_{i+1} = F_i^j$ and let F_{i+1} be $F_i^{3-x_{i+1}}$, the *other* immediate subtree of F_i . From this definition, we obtain a finite sequence $E_0, E_1, \dots, E_p, F_p$ with $E_0 = D$ and F_p a tree of height 0. Therefore, F_p consists of only one node with a label, let us say, s . Note that s is also a leaf-label of the tree D . Define now $g_B(D)$ as follows using the fact that we have g_{B_1} and g_{B_2} :

$$g_B(D) := ((g_{B_{x_1}}(E_1), \dots, g_{B_{x_p}}(E_p)), s).$$

Note that $B_{x_i} \not\leq E_i$, which means that $E_i \in L_{B_{x_i}}$. So $g_{B_{x_i}}(E_i)$ is well-defined, hence $g_B(D) \in W_B(\mathcal{T}(\mathcal{B}(\cdot)))$. Does g satisfies the desired properties?

First, we already noted that s is a leaf-label of $D(d_1, \dots, d_k)$. Secondly, if $g_{B_{x_i}}(E_i)$ is equal to $w_i(s_1^i, \dots, s_{n_i}^i)$, then by the induction hypothesis and the fact that E_i is a subtree of D , we obtain $\{s_1^i, \dots, s_{n_i}^i\} \subseteq \{d_1, \dots, d_k\}$. Hence, if

$g_B(D) = w(d'_1, \dots, d'_m) \in W_B(\mathcal{T}(\mathcal{B}(\cdot)))$, then $\{d'_1, \dots, d'_m\} \subseteq \{d_1, \dots, d_k\}$. Now we want to prove that g is a quasi-embedding. Let $\overline{E}_0, \overline{E}_1, \dots, \overline{E}_q, \overline{F}_q$ and y_1, \dots, y_q be the finite sequences forthcoming from definition (3), the definitions of x_{i+1} and F_{i+1} , but now starting with $\overline{D}(\overline{d}_1, \dots, \overline{d}_l) \in L_B$. Denote the label of the tree \overline{F}_q of height zero as \overline{s} . Assume furthermore that

$$\begin{aligned} g_B(D) &= ((g_{B_{x_1}}(E_1), \dots, g_{B_{x_p}}(E_p)), s) \\ &\leq_{W_B(\mathcal{T}(\mathcal{B}(\cdot)))} g_B(\overline{D}) = ((g_{B_{y_1}}(\overline{E}_1), \dots, g_{B_{y_q}}(\overline{E}_q)), \overline{s}). \end{aligned} \quad (4)$$

We show that inequality $D(d_1, \dots, d_k) \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} \overline{D}(\overline{d}_1, \dots, \overline{d}_l)$ holds by induction on q .

From (4), we obtain $s \leq_{\mathcal{T}(\mathcal{B}(\cdot))} \overline{s}$. Furthermore, there exist indices $1 \leq i_1 < \dots < i_p \leq q$ such that $g_{B_{x_j}}(E_j) \leq g_{B_{y_{i_j}}}(\overline{E}_{i_j})$. Because the left hand side of this inequality is in $W_{B_{x_j}}(\mathcal{T}(\mathcal{B}(\cdot)))$ and the right hand side in $W_{B_{y_{i_j}}}(\mathcal{T}(\mathcal{B}(\cdot)))$, we obtain $x_j = y_{i_j}$ for every j . Furthermore, $E_j \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} \overline{E}_{i_j}$ for every j , because $g_{B_{x_j}}$ is a quasi-embedding.

If $q = 0$, then also $p = 0$. Therefore, D is a tree of height zero with leaf-label s and \overline{D} is a tree of the same height with leaf-label \overline{s} . Hence $D \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} \overline{D}$. Now let $q > 0$. By construction,

$$\begin{aligned} g_B(F_1) &= ((g_{B_{x_2}}(E_2), \dots, g_{B_{x_p}}(E_p)), s), \\ g_B(\overline{F}_{i_1}) &= ((g_{B_{y_{i_1+1}}}(\overline{E}_{i_1+1}), \dots, g_{B_{y_q}}(\overline{E}_q)), \overline{s}). \end{aligned}$$

Because $g_B(D) \leq_{W_B(\mathcal{T}(\mathcal{B}(\cdot)))} g_B(\overline{D})$ and $g_{B_{x_1}}(E_1)$ is mapped onto $g_{B_{y_{i_1}}}(\overline{E}_{i_1})$, we obtain $g_B(F_1) \leq_{W_B(\mathcal{T}(\mathcal{B}(\cdot)))} g_B(\overline{F}_{i_1})$. Hence, by the sub-induction hypothesis on q , we have the inequality $F_1 \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} \overline{F}_{i_1}$. We also know that $E_1 \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} \overline{E}_{i_1}$ and $x_1 = y_{i_1}$. If $x_1 = 1$, then E_1 is the left-immediate subtree of $F_0 = D$ and \overline{E}_{i_1} is the left-immediate subtree of \overline{F}_{i_1-1} . Furthermore, F_1 is the right-immediate subtree of $F_0 = D$ and \overline{F}_{i_1} is the right-immediate subtree of \overline{F}_{i_1-1} . We conclude that $D = F_0 \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} \overline{F}_{i_1-1} \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} \overline{F}_0 = \overline{D}$. The same argument holds for $x_1 = 2$. Therefore, g_B is a quasi-embedding.

Assume $t = \circ[B(t_1, \dots, t_n)] \in \mathcal{T}(\mathcal{B}(\cdot))$ with $B(t_1, \dots, t_n)$ a binary tree in the partial ordering $\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))$ and assume that $L(t_i)$ are wpo's and $l(t_i) < \vartheta(\varepsilon_{\Omega+1})$. We want to prove that $L(t)$ is a wpo and $l(t) < \vartheta(\varepsilon_{\Omega+1})$.

First of all, we define a quasi-embedding f from $L(t)$ into $\mathcal{T}(W_B)$. First note that $d = \circ[D(d_1, \dots, d_k)] \in L(t)$ iff $d_i \in L(t)$ and $D(d_1, \dots, d_k) \in L_B$. Define $f(\circ)$ as \circ . Suppose $d = \circ[D(d_1, \dots, d_k)] \in L(t)$ and that $f(d_1), \dots, f(d_k)$ are already defined. If $g_B(D) = w(d'_1, \dots, d'_m) \in W_B(\mathcal{T}(\mathcal{B}(\cdot)))$ with $\{d'_1, \dots, d'_m\} \subseteq \{d_1, \dots, d_k\}$,

define $f(d)$ as $\circ[w(f(d'_1), \dots, f(d'_m))] \in \mathcal{T}(W_B)$. Now we want to prove that f is a quasi-embedding.

Assume $f(d) \leq_{\mathcal{T}(W_B)} f(\bar{d})$. We prove that this implies $d \leq_{\mathcal{T}(\mathcal{B}(\cdot))} \bar{d}$ by induction on the sum of the complexities of d and \bar{d} . If d or \bar{d} is equal to \circ , then this is trivial. Assume $g_B(D(d_1, \dots, d_k)) = w(d'_1, \dots, d'_m)$, $g_B(\bar{D}(\bar{d}_1, \dots, \bar{d}_l)) = \bar{w}(\bar{d}'_1, \dots, \bar{d}'_p)$ with $\{d'_1, \dots, d'_m\} \subseteq \{d_1, \dots, d_k\}$ and $\{\bar{d}'_1, \dots, \bar{d}'_p\} \subseteq \{\bar{d}_1, \dots, \bar{d}_l\}$ and

$$\begin{aligned} f(d) &= f(\circ[D(d_1, \dots, d_k)]) = \circ[w(f(d'_1), \dots, f(d'_m))] \\ &\leq_{\mathcal{T}(W_B)} f(\bar{d}) = f(\circ[\bar{D}(\bar{d}_1, \dots, \bar{d}_l)]) = \circ[\bar{w}(f(\bar{d}'_1), \dots, f(\bar{d}'_p))]. \end{aligned}$$

This implies either $f(d) \leq_{\mathcal{T}(W_B)} f(\bar{d}'_i)$ for some i or

$$w(f(d'_1), \dots, f(d'_m)) \leq_{W_B(\mathcal{T}(W_B))} \bar{w}(f(\bar{d}'_1), \dots, f(\bar{d}'_p)).$$

In the former case, the induction hypothesis yields $d \leq_{\mathcal{T}(\mathcal{B}(\cdot))} \bar{d}'_i \leq_{\mathcal{T}(\mathcal{B}(\cdot))} \bar{d}$. In the latter case, we observe that the induction hypothesis implies that f is a quasi-embedding from $\{d_1, \dots, d_k, \bar{d}_1, \dots, \bar{d}_l\}$ to $\mathcal{T}(W_B)$. Therefore, the Lifting Lemma brings the inequality $g_B(D) \leq_{W_B(\mathcal{T}(W_B))} g_B(\bar{D})$. Hence we have $D(d_1, \dots, d_k) \leq_{\mathcal{B}(\mathcal{T}(\mathcal{B}(\cdot)))} \bar{D}(\bar{d}_1, \dots, \bar{d}_l)$ and $d \leq_{\mathcal{T}(\mathcal{B}(\cdot))} \bar{d}$. We conclude that f is a quasi-embedding.

By Lemma 2.7 and Theorem 3.3 we obtain that $L(t)$ is a $\text{wp}\circ$ and

$$o(L(t)) \leq o(\mathcal{T}(W_B)) \leq \vartheta(W_B(\Omega)).$$

So if $\vartheta(W_B(\Omega)) < \vartheta(\varepsilon_{\Omega+1})$, we can end this proof. We prove simultaneously by induction on the height of the tree B that

$$\begin{aligned} W_B(\Omega) &< \Omega_n[1], \\ k(W_B(\Omega)) &< \vartheta(\Omega_n[1]), \end{aligned}$$

for a certain natural number n . Remark that we write $W_B(\Omega)$ instead of $o(W_B(\Omega))$ for notational convenience. From this it follows that $\vartheta(W_B(\Omega)) < \vartheta(\varepsilon_{\Omega+1})$.

If the height of the tree B is zero, we defined $W_B(X)$ as $\mathcal{B}(L_{\mathcal{T}(\mathcal{B}(\cdot))}(t_i))$. Because we assumed that $l(t_i) < \vartheta(\varepsilon_{\Omega+1})$, we know that $l(t_i) < \vartheta(\Omega_m[1]) < \Omega$ for a certain natural number $m \geq 2$. Hence

$$\begin{aligned} k(W_B(\Omega)) &= k(o(\mathcal{B}(L(t_i)))) = o(\mathcal{B}(L(t_i))) \leq \varepsilon_{l(t_i)+1} \\ &< \varepsilon_{\vartheta(\Omega_m[1])+1} = \vartheta(\Omega + \vartheta(\Omega_m[1])) \leq \vartheta(\Omega_{m+1}[1]) \end{aligned}$$

and

$$W_B(\Omega) \leq \varepsilon_{l(t_i)+1} < \Omega < \Omega_{m+1}[1].$$

Assume that the height of B is strictly larger than zero such that B_1 and B_2 are immediate subtrees of B . Because of the induction hypothesis, we know that there exists a natural number m such that

$$\begin{aligned} W_{B_1}(\Omega) &< \Omega_m[1], \\ k(W_{B_1}(\Omega)) &< \vartheta(\Omega_m[1]), \\ W_{B_2}(\Omega) &< \Omega_m[1], \\ k(W_{B_2}(\Omega)) &< \vartheta(\Omega_m[1]). \end{aligned}$$

We defined $W_B(X)$ as $(W_{B_1}(X) + W_{B_2}(X))^* \times X$, hence

$$W_B(\Omega) < \omega^{\omega^{\Omega_m[1] \oplus \Omega_m[1] \oplus 1}} \otimes \Omega < \Omega_{n_1}[1]$$

for a certain n_1 large enough. Furthermore, by Lemma 2.19

$$k(W_B(\Omega)) = k\left(\omega^{\omega^{W_{B_1}(\Omega) \oplus W_{B_2}(\Omega)(\pm 1)}} \otimes \Omega\right) < \vartheta(\Omega_{n_2}[1])$$

for some n_2 large enough. Taking $n = \max\{n_1, n_2\}$, we can end this proof. \square

Before we prove that the Howard-Bachmann ordinal is a lower bound of the maximal order type of $\mathcal{T}(\mathcal{B}(\cdot))$, we give a lemma.

Lemma 4.2. *Suppose $\alpha < \vartheta(\varepsilon_{\Omega+1})$ and $\alpha \in P$, the set of additively closed ordinals. Then α can be uniquely written as*

$$\alpha = \vartheta(\beta),$$

with $\beta = 0$ or

$$\beta = {}_{NF}\Omega^{\beta_1}\gamma_1 + \cdots + \Omega^{\beta_n}\gamma_n < \varepsilon_{\Omega+1},$$

such that $\beta > \beta_1 > \cdots > \beta_n$ and $0 < \gamma_i < \Omega$ and $k(\beta) < \alpha$.

Proof. This follows from standard properties of the ϑ -function. A proof of this fact can be found in an unpublished article of Buchholz. \square

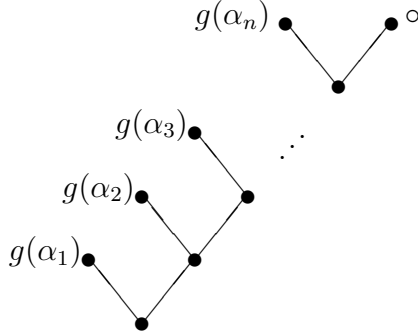
If you look closer at the proof of the next theorem, one can see how every ordinal number below $\vartheta(\varepsilon_{\Omega+1})$ can be represented as an element of $\mathcal{T}(\mathcal{B}(\cdot))$. Note that this proof can be carried out in ACA_0 if we have a predefined primitive recursive ordinal notation system for $\vartheta(\varepsilon_{\Omega+1})$.

Theorem 4.3. $o(\mathcal{T}(\mathcal{B}(\cdot))) \geq \vartheta(\varepsilon_{\Omega+1})$.

Proof. Define

$$g : \vartheta(\varepsilon_{\Omega+1}) \rightarrow \mathcal{T}(\mathcal{B}(\cdot))$$

in the following recursive way. Let $g(0)$ be \circ . Pick an arbitrary $\alpha < \vartheta(\varepsilon_{\Omega+1})$ and assume that $g(\beta)$ is already defined for every $\beta < \alpha$. If $\alpha =_{CNF} \omega^{\alpha_1} + \dots + \omega^{\alpha_n}$ with $n \geq 2$, define $g(\alpha)$ as $\circ[B]$ with B the following binary tree:

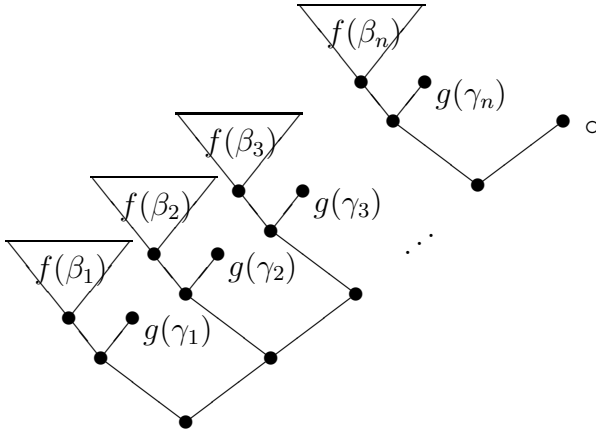


If $\alpha < \vartheta(\varepsilon_{\Omega+1})$ and $\alpha \in P$, we can write α as $\vartheta(\beta)$ as in Lemma 4.2. Because every element in $K(\beta)$ is strictly smaller than α , we can assume that $g(\gamma)$ is defined for every $\gamma \in K(\beta)$. Define $g(\alpha)$ as $\circ[f(\beta)]$, where we define the binary tree $f(\beta)$ in the following recursive way.

Let $f(0)$ be the binary tree with one node and leaf-label \circ :



Now, let $\beta = \Omega^{\beta_1}\gamma_1 + \dots + \Omega^{\beta_n}\gamma_n > \beta_1 > \dots > \beta_n \geq 0$ and assume that $f(\beta_i)$ is already defined for every i . Then define $f(\beta)$ as



Note that all labels in the tree $f(\beta)$ are elements of $g(K(\beta) \cup \{0\})$. Additionally, every element in $g(K(\beta) \cup \{0\})$ is a label in the tree $f(\beta)$.

Is g a quasi-embedding? We will show by induction on α' that $g(\alpha) \leq g(\alpha')$ implies $\alpha \leq \alpha'$. If α or α' is equal to zero, then this is trivial, hence we may assume that both α and α' are different from zero. There are now four cases left:

$$a) \ g(\omega^{\alpha_1} + \dots + \omega^{\alpha_n}) \leq g(\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m}).$$

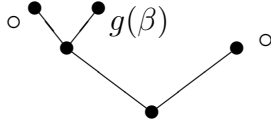
Then either $g(\alpha) \leq g(\alpha'_i)$ or $(g(\alpha_1), \dots, g(\alpha_n)) \leq^* (g(\alpha'_1), \dots, g(\alpha'_m))$. In both cases the induction hypothesis yields $\omega^{\alpha_1} + \dots + \omega^{\alpha_n} \leq \omega^{\alpha'_1} + \dots + \omega^{\alpha'_m}$.

$$b) \ g(\vartheta(\beta)) \leq g(\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m}).$$

If $\beta = 0$, then it is trivial. Assume $\beta \neq 0$. Then $g(\vartheta(\beta)) \leq g(\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m})$ is only possible if $g(\vartheta(\beta)) \leq g(\alpha'_i)$ for a certain i . The induction hypothesis yields $\vartheta(\beta) \leq \omega^{\alpha'_1} + \dots + \omega^{\alpha'_m}$.

$$c) \ g(\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m}) \leq g(\vartheta(\beta)).$$

It is impossible that $\beta = 0$, because $m \geq 2$. If $0 < \beta < \Omega$, then $g(\vartheta(\beta))$ is equal to



So $g(\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m}) \leq g(\vartheta(\beta))$ can only occur if $g(\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m}) \leq g(\beta)$ because $m \geq 2$. By the induction hypothesis, we obtain that $\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m} \leq \beta < \vartheta(\beta)$. In the case that $\beta \geq \Omega$, we have that $g(\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m}) \leq g(\gamma)$ for a certain $\gamma \in K(\beta) \cup \{0\}$ or for every α'_i , there exists a $\gamma_i \in K(\beta) \cup \{0\}$ such that $g(\alpha'_i) \leq g(\gamma_i)$. In the former case, we obtain that $\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m} \leq \gamma \leq k(\beta) < \vartheta(\beta)$. In the latter case, we obtain that $\alpha'_i \leq \gamma_i \leq k(\beta) < \vartheta(\beta)$. Because $\vartheta(\beta)$ is an epsilon number ($\beta \geq \Omega$), we get that $\omega^{\alpha'_1} + \dots + \omega^{\alpha'_m} < \vartheta(\beta)$.

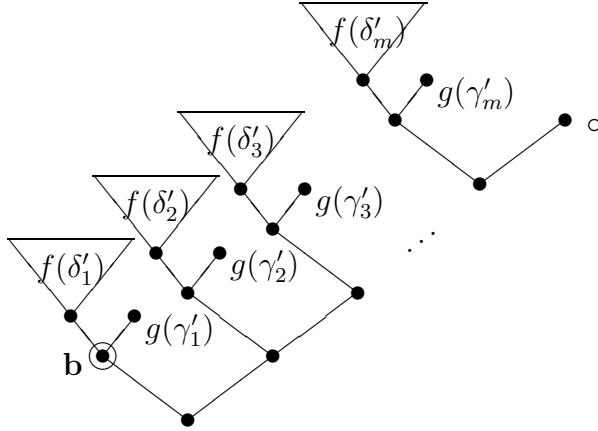
$$d) \ g(\vartheta(\beta)) \leq g(\vartheta(\beta'))$$

If $\beta' = 0$, then β must also be zero, hence $\vartheta(\beta) \leq \vartheta(\beta')$. Now assume that $\beta' > 0$. $g(\vartheta(\beta)) \leq g(\vartheta(\beta'))$ is possible if $g(\vartheta(\beta)) \leq g(\gamma')$ for a certain $\gamma' \in K(\beta') \cup \{0\}$ or if $f(\beta) \leq f(\beta')$. In the former case, we obtain by the induction hypothesis that $\vartheta(\beta) \leq \gamma' \leq k(\beta') < \vartheta(\beta')$. If the latter case occurs, then for every $\gamma \in K(\beta)$, there exists a $\gamma' \in K(\beta') \cup \{0\}$ such that $g(\gamma) \leq g(\gamma')$. Hence, $k(\beta) \leq k(\beta') < \vartheta(\beta')$. For ending the proof of $\vartheta(\beta) \leq \vartheta(\beta')$, we need to show that

$$\beta \leq \beta'.$$

We proof by induction on δ' that $f(\delta) \leq f(\delta')$ implies $\delta \leq \delta'$ for every δ with $K(\delta) \subseteq K(\beta)$ and every δ' with $K(\delta') \subseteq K(\beta')$. If this is true can conclude that $\beta \leq \beta'$.

If $\delta' = 0$ or $\delta = 0$, then this is trivial. Hence we may assume that both δ and δ' are different from zero. Assume $\delta' =_{NF} \Omega^{\delta'_1} \gamma'_1 + \dots + \Omega^{\delta'_m} \gamma'_m > 0$. Then $f(\delta')$ is equal to



There are four different subcases:

i) The root of $f(\delta)$ is mapped into $f(\delta'_1)$

Then $f(\delta) \leq f(\delta'_1)$, hence, by the induction hypothesis $\delta \leq \delta'_1 \leq \delta'$.

ii) The root of $f(\delta)$ is mapped on \mathbf{b}

This is only possible if $\delta = \Omega^{\delta_1} \gamma_1$. In this case we have that $f(\delta_1) < f(\delta'_1)$, hence $\delta_1 < \delta'_1$. This implies $\delta < \delta'$.

iii) The root of $f(\delta)$ is mapped into the right immediate subtree of $f(\delta')$

If $m = 1$, then δ has to be 0, hence $\delta \leq \delta'$. If $m > 1$, then $f(\delta) \leq f(\delta')$ yields $f(\delta) \leq f(\Omega^{\delta'_2} \gamma'_2 + \dots + \Omega^{\delta'_m} \gamma'_m)$, hence $\delta < \Omega^{\delta'_1} \gamma'_1 + \dots + \Omega^{\delta'_m} \gamma'_m = \delta'$.

iv) The root of $f(\delta)$ is mapped on the root of $f(\delta')$

Let $\delta = \Omega^{\delta_1} \gamma_1 + \dots + \Omega^{\delta_n} \gamma_n$. If the immediate left subtree of $f(\delta)$ is mapped into $f(\delta'_1)$, then like in case *ii*), we obtain that $\delta < \delta'$. So suppose not, then $f(\delta_1) \leq f(\delta'_1)$ and $g(\gamma_1) \leq g(\gamma'_1)$. Hence $\delta_1 \leq \delta'_1$ and $\gamma_1 \leq \gamma'_1$. If $\delta_1 < \delta'_1$, then $\delta < \delta'$. Assume $\delta_1 = \delta'_1$. If $\gamma_1 < \gamma'_1$, then $\delta < \delta'$. Hence assume $\gamma_1 = \gamma'_1$. If $m = 1$,

then $\delta = \delta'$. So assume $m > 1$. There are two cases:

- $n = 1$

We obtain easily that $\delta < \delta'$.

- $n \geq 2$

Then $f(\Omega^{\delta_2}\gamma_2 + \dots + \Omega^{\delta_n}\gamma_n) \leq f(\Omega^{\delta'_2}\gamma'_2 + \dots + \Omega^{\delta'_m}\gamma'_m)$. Hence the induction hypothesis yields $\Omega^{\delta_2}\gamma_2 + \dots + \Omega^{\delta_n}\gamma_n \leq \Omega^{\delta'_2}\gamma'_2 + \dots + \Omega^{\delta'_m}\gamma'_m$, hence $\delta \leq \delta'$. \square

5 Bounds on the proof-theoretical ordinals of $RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-$ and $RCA_0^* + (\Pi_1^1(\Pi_3^0) - CA_0)^-$

In the last section, we show some independence and provability results about the already studied wpo's. To obtain such results, we need bounds on the proof-theoretical ordinals of theories with light-face Π_1^1 -comprehension, which are obtained in this intermediate section. More specifically, we calculate the proof-theoretical ordinal of $RCA_0 + (\Pi_1^1 - CA_0)^-$. As a side question, we were wondering what would happen with the ordinal if we replace RCA_0 by RCA_0^* . We could not pinpoint down the exact strength of these theories, but we could do it for restricted versions. These restricted theories are fortunately strong enough to obtain the independence and provability results that we want.

Definition 5.1. Let $n \leq \omega$. A $\Pi_1^1(\Pi_n^0)$ -formula is a formula of the form $\forall X B(X)$, where $B(X)$ is Π_n^0 . B can contain set and numerical parameters. Note that a Π_ω^0 formula is the same as an arithmetical formula, hence a $\Pi_1^1(\Pi_\omega^0)$ is a standard Π_1^1 -formula. A $(\Pi_1^1(\Pi_n^0))^-$ -formula is a formula of the form $\forall X B(X)$, where $B(X)$ is Π_n^0 and $\forall X B(X)$ contains no free set parameters. It is allowed that B contains numerical parameters.

Definition 5.2. Let $(\mathcal{F} - CA_0)$ be the following well-known comprehension scheme

$$\exists Z \forall n (n \in Z \leftrightarrow A(n)),$$

where $A(n)$ is a formula in the class \mathcal{F} . If $\mathcal{F} = (\Pi_1^1(\Pi_n^0))^-$, we denote $(\mathcal{F} - CA_0)$ by $(\Pi_1^1(\Pi_n^0) - CA_0)^-$

We show the following results.

Theorem 5.3.

1. $|RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-| = \vartheta(\Omega^\omega)$,
2. $|RCA_0^* + (\Pi_1^1(\Pi_3^0) - CA_0)^-| \geq \varphi\omega 0 = \vartheta(\Omega \cdot \omega)$.

For clarity, we give the definition of ACA_0 , RCA_0 and RCA_0^*

Definition 5.4. *Both ACA_0 and RCA_0 are theories in the language of second order arithmetic L_2 . The axioms of ACA_0 are the basic axioms in second order arithmetic and the induction axiom for sets together with comprehension for arithmetical formulas. RCA_0 consists of the basic axioms, Σ_1^0 -induction and Δ_1^0 -comprehension. Let $L_2(\exp)$ be the language L_2 augmented by a binary operation symbol \exp that denotes the exponential function. Define RCA_0^* as the $L_2(\exp)$ -theory consisting of the basic axioms, the exponentiation axioms, Σ_0^0 -induction and Δ_1^0 -comprehension.*

From more information on reverse mathematics and theories in second order arithmetic, we refer the reader to [22].

5.1 Lower bounds

These proofs and definitions follow the procedure as in [17], but they need some refinements. Firstly, we give a primitive recursive ordinal notation system for $\vartheta(\varepsilon_{\Omega+1})$ which is suitable for using in $ACA_0 + (\mathcal{F}-CA_0)$. Then, we introduce an notation system without ω -exponentiation which is suitable to use in $RCA_0 + (\mathcal{F}-CA_0)$ and $RCA_0^* + (\mathcal{F}-CA_0)$.

Definition 5.5. *Define inductively a set $OT(\vartheta)$ of ordinals and a natural number $G_\vartheta\alpha$ for $\alpha \in OT(\vartheta)$ as follows:*

1. $0 \in OT(\vartheta)$ and $G_\vartheta(0) := 0$,
2. if $\alpha = \Omega^{\alpha_1}\beta_1 + \dots + \Omega^{\alpha_n}\beta_n$ with $n \geq 1$, $\alpha_1 > \dots > \alpha_n$ and $\Omega > \beta_1, \dots, \beta_n > 0$, then
 - (a) if $(n > 1 \text{ or } \alpha_1 > 0)$ and $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in OT(\vartheta)$, then $\alpha \in OT(\vartheta)$ and $G_\vartheta\alpha := \max\{G_\vartheta(\alpha_1), \dots, G_\vartheta(\alpha_n), G_\vartheta(\beta_1), \dots, G_\vartheta(\beta_n)\} + 1$,
 - (b) if $n = 1$, $\alpha_1 = 0$ and $\alpha = \omega^{\delta_1} + \dots + \omega^{\delta_m} > \delta_1 \geq \dots \geq \delta_m$ with $m \geq 2$ and $\delta_1, \dots, \delta_m \in OT(\vartheta)$, then $\alpha \in OT(\vartheta)$ and $G_\vartheta\alpha := \max\{G_\vartheta(\delta_1), \dots, G_\vartheta(\delta_m)\} + 1$,
3. if $\alpha = \vartheta\beta$ and $\beta \in OT(\vartheta)$, then $\alpha \in OT(\vartheta)$ and $G_\vartheta\alpha := G_\vartheta\beta + 1$.

Because $\vartheta\beta$ is always additively closed and ϑ is injective, the function G_ϑ is well-defined. Remark that $\Omega \in OT(\vartheta)$, because $\Omega = \Omega^1 \cdot 1$ and $1 = \vartheta(0)$.

Lemma 5.6. *If $\xi \in OT(\vartheta)$, then $K(\xi) \subseteq OT(\vartheta)$. Furthermore, $G_\vartheta(k(\xi)) \leq G_\vartheta(\xi)$ for all ξ in $OT(\vartheta)$.*

Proof. We proof this by induction on $G_\vartheta(\xi)$. If $\xi = 0$, then this trivially holds. If $\xi = \vartheta(\xi')$, then $K(\xi) = \{\xi\}$, hence this also trivially holds. Assume $\xi = \Omega^{\xi_1}\beta_1 + \dots + \Omega^{\xi_n}\beta_n$ with $n \geq 1$, $\xi_1 > \dots > \xi_n$ and $\Omega > \xi_1, \dots, \xi_n > 0$. If $n = 1$ and $\xi_1 = 0$, then also $K(\xi) = \{\xi\}$, hence the proof is valid. Let now $n > 1$ or $\xi_1 > 0$. Then $K(\xi) = \{\beta_1, \dots, \beta_n\} \cup K(\xi_1) \cup \dots \cup K(\xi_n)$. The induction hypothesis yields $K(\xi) \subseteq OT(\vartheta)$. Furthermore, $G_\vartheta(k(\xi_i)) \leq G_\vartheta(\xi_i) < G_\vartheta(\xi)$. Therefore, the strict inequality $G_\vartheta(k(\xi)) = G_\vartheta(\max_i\{k(\xi_i), \beta_i\}) < G_\vartheta(\xi)$ holds. \square

Each ordinal $\alpha \in OT(\vartheta)$ has a unique normal form using the symbols $0, \omega, \Omega, +, \vartheta$. The relation $\alpha < \beta$ can be expressed using the ordinals appearing in their normal form (by Lemma 2.18), which have strictly smaller G_ϑ -values by the previous Lemma 5.6. Hence, the following lemma holds.

Lemma 5.7. *If we use a specific coding of $(OT(\vartheta), <_{OT(\vartheta)})$ in the natural numbers, then $(OT(\vartheta) \cap \Omega, <_{OT(\vartheta)})$ can be interpreted as a primitive recursive ordinal notation system for the ordinal $\vartheta(\varepsilon_{\Omega+1})$. Furthermore, one can choose this coding in such a way that $\forall \xi \in K(\alpha)(\xi \leq_N \alpha)$.*

This is the ordinal notation system that we will use if we work in the theory $ACA_0 + (\mathcal{F}-CA_0)$. However, in the theory $RCA_0 + (\mathcal{F}-CA_0)$ and $RCA_0^* + (\mathcal{F}-CA_0)$, we use a different ordinal notation system $OT'(\vartheta)$ without ω -exponentiation.

Definition 5.8. *Define inductively a set $OT'(\vartheta)$ of ordinals and a natural number $G'_\vartheta\alpha$ for $\alpha \in OT'(\vartheta)$ as follows:*

1. $0 \in OT'(\vartheta)$ and $G'_\vartheta(0) := 0$,
2. if $\alpha = \Omega^n\beta_n + \dots + \Omega^0\beta_0$, with $\Omega > \beta_n > 0$ and $\Omega > \beta_0, \dots, \beta_{n-1}$, then
 - (a) if $n > 0$ and $\beta_0, \dots, \beta_n \in OT'(\vartheta)$, then $\alpha \in OT'(\vartheta)$ and $G'_\vartheta\alpha := \max\{G'_\vartheta(\beta_0), \dots, G'_\vartheta(\beta_n)\} + 1$,
 - (b) if $n = 0$ and $\alpha =_{NF} \delta_1 + \dots + \delta_m$ with $m \geq 2$ and $\delta_1, \dots, \delta_m \in OT'(\vartheta)$, then $\alpha \in OT'(\vartheta)$ and $G'_\vartheta\alpha := \max\{G'_\vartheta(\delta_1), \dots, G'_\vartheta(\delta_m)\} + 1$,
3. if $\alpha = \vartheta\beta$ and $\beta \in OT'(\vartheta)$, then $\alpha \in OT'(\vartheta)$ and $G'_\vartheta\alpha := G'_\vartheta\beta + 1$.

NF stands for normal form.

Also in this ordinal notation system, G'_ϑ is well-defined.

Lemma 5.9. *If $\xi \in OT'(\vartheta)$, then $K(\xi) \subseteq OT'(\vartheta)$. Furthermore, $G'_\vartheta(k(\xi)) \leq G'_\vartheta(\xi)$ for all ξ in $OT'(\vartheta)$.*

Proof. If $\xi = 0$, then this trivially holds. If $\xi = \vartheta(\xi')$, then $K(\xi) = \{\xi\}$, hence this also trivially holds. Assume $\xi = \Omega^n \xi_n + \dots + \Omega^0 \xi_0$ with $\Omega > \xi_n > 0$ and $\Omega > \xi_1, \dots, \xi_{-1}$. If $n = 0$, then also $K(\xi) = \{\xi\}$, hence the proof is valid. Let now $n > 0$. Then $K(\xi) \subseteq \{n, \dots, 0, \xi_0, \dots, \xi_n\}$. We know $\xi_0, \dots, \xi_n \in OT'(\vartheta)$. Additionally, it is trivial to show $\{n, \dots, 0\} \subseteq OT'(\vartheta)$. Hence, $K(\xi) \subseteq OT'(\vartheta)$. It is also trivial to show that $G'_\vartheta(m) \leq 2$ for all natural numbers m . Therefore, $G'_\vartheta(n) \leq G'_\vartheta(\xi)$ because $G'_\vartheta(\xi_n) \geq 1$. Also, $G'_\vartheta(\xi_i) < G'_\vartheta(\xi)$. Hence $G'_\vartheta(k(\xi)) = G'_\vartheta(\max_i \{n, \xi_i\}) < G'_\vartheta(\xi)$. \square

Like for the first ordinal notation system, we have to following lemma.

Lemma 5.10. *If we use a specific coding of $(OT'(\vartheta), <_{OT'(\vartheta)})$ in the natural numbers, $(OT'(\vartheta) \cap \Omega, <_{OT'(\vartheta)})$ can be interpreted as a primitive recursive ordinal notation system for the ordinal $\vartheta(\Omega^\omega)$. Furthermore, one can choose this coding in such a way that $\forall \xi \in K(\alpha)(\xi \leq_N \alpha)$.*

From now on, we fix primitive recursive ordinal notation systems $OT(\vartheta) \cap \Omega$ and $OT'(\vartheta) \cap \Omega$. If we mention ACA_0 in the beginning of a theorem, then we assume that we work in $OT(\vartheta)$. Similarly, if we mention RCA_0 or RCA_0^* in the beginning of a theorem, we assume that we work in $OT'(\vartheta)$.

Definition 5.11.

$$\begin{aligned} Prog(\prec, F) &:= \forall x (\forall y (y \prec x \rightarrow F(y)) \rightarrow F(x)), \\ TI(\prec, F) &:= Prog(\prec, F) \rightarrow \forall x F(x), \\ WF(\prec) &:= \forall X TI(\prec, X), \end{aligned}$$

where the formula $Prog(\prec, F)$ stands for progressiveness, $TI(\prec, F)$ for transfinite induction and $WF(\prec)$ for well-foundedness. $F(x)$ is an arbitrary $L_2(exp)$ -formula if we work in RCA_0^* or an arbitrary L_2 -formula if we work in RCA_0 or ACA_0 . For an element $\alpha \in OT(\vartheta)$, the formula $WF(\alpha)$ stands for ' $<_{OT(\vartheta)}$ restricted to $\{\beta \in OT(\vartheta) : \beta < \alpha\}$ is well-founded'. We sometimes also denote this as $WF(< \upharpoonright \alpha)$. Similarly for $OT'(\vartheta)$.

Definition 5.12. 1. $Acc := \{\alpha < \Omega : WF(< \upharpoonright \alpha)\},$

2. $M := \{\alpha : K(\alpha) \subseteq Acc\},$

3. $\alpha <_\Omega \beta \Leftrightarrow \alpha, \beta \in M \wedge \alpha < \beta.$

The next lemma shows that Acc , M and $<_\Omega$ can be expressed by a $(\Pi_1^1(\Pi_3^0))^-$ -formula.

Lemma 5.13. *Acc, M and $<_\Omega$ are expressible by a $(\Pi_1^1(\Pi_3^0))^-$ -formula.*

Proof. The proof is the same for the ordinal notation system $OT(\vartheta)$ and $OT'(\vartheta)$.

$$WF(\alpha) = \forall X (\forall x (\forall y (y \prec x \rightarrow y \in X) \rightarrow x \in X) \rightarrow \forall x (x \in X)),$$

where \prec is $< \upharpoonright \alpha$. It is easy to see that the prenex normal form of the formula $WF(\alpha)$ is $(\Pi_1^1(\Pi_3^0))^-$, hence Acc can be expressed by such a formula. M can be represented by the formula $\forall \xi \leq_N \alpha (\xi \in K(\alpha) \rightarrow \xi \in Acc)$. Because $\xi \in K(\alpha)$ is elementary recursive, both M and $<_\Omega$ are also expressible by a $(\Pi_1^1(\Pi_3^0))^-$ -formula. \square

Lemma 5.14. $(RCA_0^*) \alpha, \beta \in Acc \implies \alpha + \beta \in Acc$.

Proof. Obvious. \square

Lemma 5.15. $(ACA_0) \alpha \in Acc \implies \omega^\alpha \in Acc$.

Proof. A proof of this lemma goes back to Gentzen. See [15, 19]. \square

Definition 5.16. *Let $Prog_\Omega(X)$ be the formula*

$$(\forall \alpha \in M) [(\forall \beta <_\Omega \alpha)(\beta \in X) \rightarrow \alpha \in X].$$

Let Acc_Ω be the set $\{\alpha \in M : \vartheta(\alpha) \in Acc\}$.

Lemma 5.17. $(RCA_0^*) Prog_\Omega(Acc_\Omega)$.

Proof. We work in $OT'(\vartheta)$. Assume $\alpha \in M$ and $(\forall \beta <_\Omega \alpha)(\vartheta(\beta) \in Acc)$. We want to proof that $\vartheta(\alpha) \in Acc$. We show that $(\forall \xi < \vartheta(\alpha))(\xi \in Acc)$ by induction on $G'_\vartheta \xi$, from which the theorem follows. If $\xi = 0$, then this trivially holds. So assume $\xi > 0$.

a) Assume $\xi \notin P$

Because $\xi < \Omega$, we have $\xi =_{NF} \xi_1 + \dots + \xi_n > \xi_1 \geq \dots \geq \xi_n$ ($n \geq 2$). The induction hypothesis yields $\xi_i \in Acc$. Hence from Lemma 5.14, we obtain $\xi \in Acc$.

b) Assume $\xi = \vartheta(\xi')$

From $\vartheta(\xi') < \vartheta(\alpha)$, we obtain either $\xi' < \alpha$ and $k(\xi') < \vartheta(\alpha)$ or $\vartheta(\xi') \leq k(\alpha)$. In the former case, $G'_\vartheta(k(\xi')) \leq G'_\vartheta(\xi') < G'_\vartheta(\xi)$ and the induction hypothesis implies $k(\xi') \in Acc$, hence $K(\xi') \subseteq Acc$. So $\xi' \in M$, from which it follows $\xi' <_\Omega \alpha$, hence $\vartheta(\xi') \in Acc$. In the latter case, we know that $k(\alpha) \in Acc$, because $\alpha \in M$. Therefore, $\vartheta(\xi') \in Acc$. \square

We actually do not need the following lemma, because we already know the proof-theoretical ordinal of $ACA_0 + (\Pi_1^1 - CA_0)^-$. However, just for the completeness, we mention it here.

Lemma 5.18. $(ACA_0) \text{ Prog}_\Omega(Acc_\Omega)$.

Proof. The proof uses $OT(\vartheta)$ and follows the same procedure as Lemma 5.17. The only difference is the usage of Lemma 5.15. \square

Lemma 5.19. Let $A(a)$ be a $(\Pi_1^1(\Pi_3^0))^-$ -formula. Define A_k as

$$\forall \alpha [(\forall \beta <_\Omega \alpha) A(\beta) \rightarrow (\forall \beta <_\Omega \alpha + \Omega^k) A(\beta)].$$

Then $RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-$ proves $\text{Prog}_\Omega(\{\xi : A(\xi)\}) \rightarrow A_k$.

Proof. We will prove that by outer induction on k . First note that $A(a)$ can be expressed by a set in $RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-$. Assume that $\text{Prog}_\Omega(\{\xi : A(\xi)\})$ and pick an arbitrary α such that $(\forall \beta <_\Omega \alpha) A(\beta)$. If $\alpha \notin M$, the assertion trivially follows. Assume $\alpha \in M$. If $k = 0$, the proofs follows easily from $\text{Prog}_\Omega(\{\xi : A(\xi)\})$. Assume $k = l + 1$ en suppose $(\forall \beta <_\Omega \alpha) A(\beta)$. We want to prove that $(\forall \beta <_\Omega \alpha + \Omega^{l+1}) A(\beta)$. Take an arbitrary $\beta <_\Omega \alpha + \Omega^{l+1}$. RCA_0 proves that there exists a $\xi \in Acc$ such that $\beta <_\Omega \alpha + \Omega^l \xi$ (by induction on the construction of β in $OT(\vartheta)$, one can show that one can take $\xi = k(\beta) + 1$). Let $B(\zeta)$ be

$$(\forall \beta <_\Omega \alpha + \Omega^l \zeta) A(\beta).$$

$B(\zeta)$ is a Π_1^0 -formula in A . It is known (Lemma 6 in [9]) that $RCA_0 \vdash \forall X (WO(X) \rightarrow TI_X(\psi))$ for all Π_1^0 -formulas ψ . Because $\xi \in Acc$, we have that $\xi + 1$ is well-ordered, hence we know that $TI_{\xi+1}(B)$ is true. This means

$$\forall x \leq \xi [\forall y \leq \xi (y < x \rightarrow B(y)) \rightarrow B(x)] \rightarrow \forall x \leq \xi B(x).$$

The theorem follows from $B(\xi)$, hence we only have to prove the progressiveness of B along $\xi + 1$.

Assume that $x \leq \xi$. Then $x \in Acc$. If $x = 0$, then $B(0)$ follows from the assumption. Assume that x is a limit. If $\beta <_\Omega \alpha + \Omega^l x$, then there exists a $y < x$ such that $\beta <_\Omega \alpha + \Omega^l y$. Because $B(y)$ is valid, one obtains $A(\beta)$. Assume that $x = x' + 1 \in Acc$. From $x' < x$, one obtains $(\forall \beta <_\Omega \alpha + \Omega^l x') A(\beta)$. Because A_l is valid, we get $(\forall \beta <_\Omega \alpha + \Omega^l x' + \Omega^l) A(\beta)$, hence $(\forall \beta <_\Omega \alpha + \Omega^l (x' + 1)) A(\beta) = B(x)$. \square

Theorem 5.20. $|RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-| \geq \vartheta(\Omega^\omega)$.

Proof. This follows from Lemmas 5.17, 5.19 and the fact that Acc_Ω is expressible by a $(\Pi_1^1(\Pi_3^0))^-$ -formula. \square

Lemma 5.21. *Let $A(a)$ be a $(\Pi_1^1(\Pi_3^0))^-$ -formula. Define A_k as*

$$(\forall \beta <_\Omega \Omega \cdot k) A(\beta).$$

Then $RCA_0^ + (\Pi_1^1(\Pi_3^0) - CA_0)^-$ proves $Prog_\Omega(\{\xi : A(\xi)\}) \rightarrow A_k$.*

Proof. We will prove that by outer induction on k . First note that $A(a)$ can be expressed by a set in $RCA_0^* + (\Pi_1^1(\Pi_3^0) - CA_0)^-$. It is easy to see that the case $k = 0$ holds. Assume $k = l + 1$ and $Prog_\Omega(\{\xi : A(\xi)\})$. Then we know $(\forall \beta <_\Omega \Omega \cdot l) A(\beta)$. Pick an arbitrary $\beta <_\Omega \Omega \cdot (l + 1)$. If $\beta < \Omega \cdot l$, we obtain $A(\beta)$. Hence assume that $\beta \geq \Omega \cdot l$. There exists a $\xi < \Omega$ such that $\beta = \Omega \cdot l + \xi$. Let $B(\zeta)$ be $A(\Omega \cdot l + \zeta)$. $B(\zeta)$ is a Π_0^0 -formula in A . We prove by induction on ζ that $(\forall \zeta \in Acc) B(\zeta)$ is true. From this, the theorem follows.

If $\zeta = 0$, then $B(\zeta)$ is true because $(\forall \beta <_\Omega \Omega \cdot l) A(\beta)$ and $Prog_\Omega(\{\xi : A(\xi)\})$ imply $A(\Omega \cdot l)$. Assume $\zeta \in Acc$ is a limit and assume $(\forall \zeta' <_\Omega \zeta) B(\zeta')$. From $Prog_\Omega(\{\xi : A(\xi)\})$, we obtain $B(\zeta)$. Let $\zeta = \zeta' + 1 \in Acc$. Then $B(\zeta)$ follows from $B(\zeta')$ and $Prog_\Omega(\{\xi : A(\xi)\})$. \square

Theorem 5.22. $|RCA_0^* + (\Pi_1^1(\Pi_3^0) - CA_0)^-| \geq \vartheta(\Omega \cdot \omega) = \varphi\omega 0$.

Proof. This follows from Lemmas 5.17, 5.21 and the fact that Acc_Ω is expressible by a $(\Pi_1^1(\Pi_3^0))^-$ -formula. \square

5.2 Upper bounds

In this subsection, we give an upper bound for $|RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-|$. For this, we use the fact that $|\Pi_2^1 - BI_0| = \vartheta(\Omega^\omega)$ (see [17]).

Lemma 5.23. *For every arithmetical formula $B(X)$ with all free set variables indicated, there is a Δ_0 -formula $R(x, X, f)$ such that*

1. $ACA_0 \vdash B(X) \rightarrow \exists f \forall x R(x, X, f)$
2. If T is a theory with $RCA_0 \subseteq T$ and $\mathcal{F}(x, y)$ is an arbitrary formula,

$$T \vdash \forall x \exists ! y \mathcal{F}(x, y) \wedge \forall x \exists z (lh(z) = x \wedge \forall i < x \mathcal{F}(i, (z)_i)),$$

then $T \vdash \forall x R(x, X, \mathcal{F}) \rightarrow B(X)$, where $R(x, X, \mathcal{F})$ results from $R(x, X, f)$ by replacing subformulae of the form $f(t) = s$ by $\mathcal{F}(t, s)$.

Proof. This proof is a little adaptation of the normal form theorem V.1.4 in [22]. We can assume that $B(X)$ is in prenex normal form

$$B(X) \equiv \forall x_1 \exists y_1 \dots \forall x_r \exists y_r S(x_1, y_1, \dots, x_r, y_r, X),$$

with S quantifier-free. Then over ACA_0 , we have

$$\begin{aligned} B(X) &\leftrightarrow \exists f_1 \forall x_1 x_2 \exists y_2 \dots S(x_1, f_1(x_1), x_2, y_2, \dots, x_r, y_r, X) \\ &\leftrightarrow \exists f_1 \exists f_2 \forall x_1 x_2 x_3 \exists y_3 \dots S(x_1, f_1(x_1), x_2, f_2(x_1, x_2), x_3, y_3, \dots, x_r, y_r, X) \\ &\dots \\ &\leftrightarrow \exists f_1 \dots f_r \forall x_1 \dots x_r S(x_1, f_1(x_1), x_2, f_2(x_1, x_2), \dots, x_r, f_r(x_1, \dots, x_r), X) \\ &\leftrightarrow \exists f \forall x S((x)_1, (f)_1((x)_1), \dots, (x)_r, (f)_r((x)_1, \dots, (x)_r), X), \end{aligned}$$

where $(f)_k((x)_1, \dots, (x)_k) = f(\langle k, (x)_1, \dots, (x)_k \rangle)$ and $\langle \dots \rangle$ is some primitive recursive tuple coding. Now let $R(x, X, f) := S((x)_1, (f)_1((x)_1), \dots, X)$. Note that the right-to-left directions are provable in RCA_0 , so that part 2. follows from reading the equivalences from bottom to top. \square

By adapting the previous lemma, we get the following corollary.

Corollary 5.24. *If in addition to the conditions of Lemma 5.23(2.), T also satisfies*

$$T \vdash \forall x \exists! y \mathcal{G}(x, y) \wedge \forall x \exists z (lh(z) = x \wedge \forall i < x \mathcal{G}(i, (z)_i)),$$

then $T \vdash \forall x R(x, \mathcal{G}, \mathcal{F}) \rightarrow B(\mathcal{G})$, where $R(x, \mathcal{G}, \mathcal{F})$ results from $R(x, X, \mathcal{F})$ by replacing $t \in X$ by $\mathcal{G}(t, 0)$.

Lemma 5.25. *Assuming that the conditions for \mathcal{F} and \mathcal{G} from Lemma 5.23 and Corollary 5.24 are satisfied, then there exists a Δ_1^0 -formula P such that*

$$T \vdash \forall x R(x, \mathcal{G}, \mathcal{F}) \leftrightarrow \forall x P(\mathcal{G}[x], \mathcal{F}[x])$$

Proof. By induction on the build-up of R . \square

Theorem 5.26. *Assume that $B(X, n)$ is a Π_3^0 -formula, then*

$$\Pi_2^1 - BI_0 \vdash \forall n [\forall A \in Rec((\Pi_1^1)^-) B(A(\cdot), n) \leftrightarrow (\forall X \subseteq \omega) B(X, n)],$$

where $B(A(\cdot), n)$ results from $B(X, n)$ by replacing $t \in X$ by $A(t)$.

Proof. Pick an arbitrary natural number n_0 . Assume that we have

$$(\exists A \in Rec((\Pi_1^1)^-)) \neg B(A(\cdot), n_0).$$

In [10] it is proven that Σ_4^1 - ω -model reflection holds in Π_2^1 - BI_0 . Therefore, there exists an ω -model M such that $n_0 \in M$ and $M \models ACA_0 + \neg B(A, n_0)$. Define X as the set $\{m \in \omega : M \models A(m)\}$. This set exists by arithmetical comprehension. Hence, $\neg B(X, n_0)$ is valid.

For the reverse implication, assume $(\exists X \subseteq \omega) \neg B(X, n_0)$. Using Lemmas 5.23 and 5.25 on the formula $\neg B(X, n_0)$, we have $\exists X \exists f \forall x P(X[x], f[x])$, thus $\exists h \forall x P(h_0[x], h_1[x])$. Define $I(\sigma)$ as $\exists h(\sigma \subseteq h \wedge \forall x P(h_0[x], h_1[x]))$. Define a class function \mathcal{H} by $\mathcal{H}(0) = \langle \rangle$ and $\mathcal{H}(n+1) = \mathcal{H}(n) \smallfrown \langle \mu k \cdot I(\mathcal{H}(n) \smallfrown \langle k \rangle) \rangle$. Note that \mathcal{H} has a Σ_2^1 -graph. As Π_2^1 - BI_0 proves Σ_2^1 -induction on the naturals, one can prove that $\mathcal{H}(n)$ is always defined by $\exists h \forall x P(h_0[x], h_1[x])$. Also note that \mathcal{H} is recursive in I . We can show that I is recursive in a $(\Pi_1^1)^-$ -formula. This because I is recursive in W , the set of all recursive well-founded trees, which is $(\Pi_1^1)^-$ -complete. Therefore, we obtain that $\mathcal{H} \in Rec((\Pi_1^1)^-)$. We also have $\forall x P(\mathcal{H}_0[x], \mathcal{H}_1[x])$ by this construction. Letting $\mathcal{G} = \mathcal{H}_0$ and $\mathcal{F} = \mathcal{H}_1$ in Lemma 5.25, we get $\forall x R(x, \mathcal{G}, \mathcal{F})$. By Corollary 5.24 we get $\neg B(\mathcal{G}, n_0)$. \square

Corollary 5.27. $|RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-| = \vartheta(\Omega^\omega)$.

Proof. Follows from Theorems 5.20, 5.26 and the fact that $|\Pi_2^1 - BI_0| = \vartheta(\Omega^\omega)$. \square

In a similar way, one could also show the following lemma.

Lemma 5.28. Assume that $B(X, n)$ is a Π_2^0 -formula, then

$$\Pi_1^1 - BI_0 \vdash \forall n [\forall A \in Rec((\Pi_1^1)^-) B(A(\cdot), n) \leftrightarrow (\forall X \subseteq \omega) B(X, n)].$$

Therefore, $|\Pi_1^1 - BI_0| = \varphi\omega 0$ yields that the ordinal of $RCA_0^* + (\Pi_1^1(\Pi_2^0) - CA_0)^-$ is bounded above by $\varphi\omega 0$. Actually one can prove that the ordinal of the theory $RCA_0^* + (\Pi_1^1(\Pi_2^0) - CA_0)^-$ is even much lower. Using WKL_0 and the normal form theorem II.2.7 in [22], one can prove that every $(\Pi_1^1(\Pi_2^0))^-$ -formula is equivalent with a $(\Pi_2^0)^-$ -formula (and if the original formula has extra parameters, then the equivalent one will also have those parameters). The intuitive idea behind this is that WKL_0 proves that the projection of closed set is a closed set, meaning that $\exists X \forall x \dots$ can be reduced to $\forall z \dots$ (a closed set can be seen in some sense as a Π_1^0 -formula). Furthermore, one can proof the following lemmas.

Lemma 5.29. Let \mathcal{F} be $(\Pi_1^1(\Pi_n^0))^-$ or $(\Pi_n^0)^-$. Then the first order part of the theory $WKL_0 + \mathcal{F} - CA_0$ is the same as that of $RCA_0 + \mathcal{F} - CA_0$.

Proof. Can be proved easily by little adaptations of paragraph IX.2 in [22]. \square

In a similar way,

Lemma 5.30. *The first order part of $RCA_0 + (\Pi_2^0)^- - IND$ is IS_3 .*

Proof. Can be proved by adaptations of paragraph IX.1 in [22]. □

Hence, $|RCA_0 + (\Pi_1^1(\Pi_2^0) - CA_0)^-| = |IS_3| = \omega^{\omega^{\omega^{\omega}}}$. In a similar but more technical way, one could prove that $RCA_0^* + (\Pi_1^1(\Pi_2^0))^-$ is Π_4 -conservative over IS_2 . Hence, $|RCA_0^* + (\Pi_1^1(\Pi_2^0))^-| = \omega^{\omega^{\omega}}$. We want to thank Leszek Kołodziejczyk, who reminded us of these facts.

There are still some open questions left, for example what is the ordinal of $RCA_0 + (\Pi_1^1(\Pi_4^0) - CA_0)^-$ or $RCA_0 + (\Pi_1^1 - CA_0)^-$ etc. We state the following conjectures

Conjecture 5.31.

1. $|RCA_0^* + (\Pi_1^1(\Pi_3^0) - CA_0)^-| = \varphi\omega 0$,
2. $|RCA_0^* + (\Pi_1^1 - CA_0)^-| = \varphi\omega 0$,
3. $|RCA_0 + (\Pi_1^1 - CA_0)^-| = \vartheta(\Omega^\omega)$.

6 Independence results

6.1 Independence results for $ACA_0 + (\Pi_1^1 - CA_0)^-$

It is well-known that $|ACA_0 + (\Pi_1^1 - CA_0)^-| = \vartheta(\varepsilon_{\Omega+1})$ (for example see [2]). Now,

Theorem 6.1. $ACA_0 \vdash \text{‘}\mathcal{T}(\mathcal{B}(\cdot)) \text{ is a wpo’} \rightarrow WF(\vartheta(\varepsilon_{\Omega+1}))$.

Proof. Theorem 4.3 can be carried out in ACA_0 if we use the primitive recursive ordinal notation system $OT(\vartheta) \cap \Omega$ for $\vartheta(\varepsilon_{\Omega+1})$. □

Therefore, one has the following theorem,

Theorem 6.2. $ACA_0 + (\Pi_1^1 - CA_0)^- \not\vdash \text{‘}\mathcal{T}(\mathcal{B}(\cdot)) \text{ is a wpo’}$.

The $\text{wpo } \mathcal{T}(\mathcal{B}(\cdot))$ can be seen as the limit of $\mathcal{T}(\overbrace{.* \cdots *}^n)$, because a binary tree can be seen as an iteration of the $*$ -operator. For example, one can interpret an element of $\{a\}^{**}$ as a binary tree which goes only one time to the left and every node has label a . So, it would be interesting if $ACA_0 + (\Pi_1^1 - CA_0)^- \vdash \text{‘}\mathcal{T}(\overbrace{.* \cdots *}^n)$

is a wpo' for every natural number n because this theory does not prove the ‘limit’. The maximal order type of $\mathcal{T}(\cdot^{*\cdots*})$ is strictly below the proof-theoretical ordinal of $ACA_0 + (\Pi_1^1 - CA_0)^-$. Therefore, one can really expect this provability assertion. However, in the proof of the upper bound of the maximal order type of $\mathcal{T}(\cdot^{*\cdots*})$ one uses several induction schemes and it is not immediately clear that the proof goes through in $ACA_0 + (\Pi_1^1 - CA_0)^-$. Theorem 6.4 shows that this is indeed possible and uses the so-called minimal bad sequence argument. For more information about the minimal bad sequence argument and its reverse mathematical strength, we refer the reader to [14].

Lemma 6.3. *Over RCA_0 , the following are equivalent*

1. ACA_0 ,
2. *Higman’s theorem, i.e. $\forall X (X \text{ is a wpo} \rightarrow X^* \text{ is a wpo})$.*
3. $\forall X (X \text{ is a well-quasi-order} \rightarrow X^* \text{ is a well-quasi-order})$.

Proof. It is trivial to show that (2.) and (3.) are equivalent. In [6] and [8], the reader finds a proof of the fact that ACA_0 is equivalent over RCA_0 with $\forall \alpha (\alpha \text{ is a well-order} \rightarrow 2^\alpha \text{ is a well-order})$. One can prove that the latter is equivalent with Higman’s theorem. For a detailed version see [5]. □

Theorem 6.4. *For all n , $ACA_0 + (\Pi_1^1 - CA_0)^- \vdash \mathcal{T}(\cdot^{\overbrace{*\cdots*}^n})$ is a wpo' .*

Proof. Fix a natural number n . We reason in $ACA_0 + (\Pi_1^1 - CA_0)^-$. Code the elements of $\mathcal{T}(\cdot^{\overbrace{*\cdots*}^n})$ as natural numbers such that \circ has code 0 and the code of t_i is strictly smaller than the code of $t = \circ[w(t_1, \dots, t_i, \dots, t_n)]$. This coding can be done primitive recursively. We say that the leaves of $\circ[w(t_1, \dots, t_i, \dots, t_n)]$ are $\{t_1, \dots, t_n\}$ and we assume that there is a primitive recursive relation ‘... is a leaf of ...’. If σ decodes a finite sequence, then this sequence is equal to $((\sigma)_0, \dots, (\sigma)_{lh(\sigma)-1})$, where $lh(\sigma)$ is the length and $(\sigma)_i$ is the i -th element of the sequence. An infinite sequence $(\sigma_i)_{i < \omega}$ is decoded as a set $\{(i, \sigma_i) : i < \omega\}$. $(\{(i, \sigma_i) : i < \omega\})_i$ stands for the element σ_i . Note that one can recursively construct the set $\{(\sigma_0, \dots, \sigma_i) : i < \omega\}$ from the original set $\{(i, \sigma_i) : i < \omega\}$, where $(\sigma_0, \dots, \sigma_i)$ is decoded as a natural number. Furthermore, if one has the set $\{(\sigma_0, \dots, \sigma_i) : i < \omega\}$, one can reconstruct the original set $\{(i, \sigma_i) : i < \omega\}$ from it in a recursive way.

Now, assume that $\mathcal{T}(\cdot^{***})$ is not a well-partial-order. Then there exists an infinite sequence $(t_i)_{i < \omega}$ in $\mathcal{T}(\cdot^{***})$ such that $\forall i, j (i < j \rightarrow t_i \not\leq t_j)$. Define $\chi(\sigma)$ as

σ is a finite sequence of elements in $\mathcal{T}(\cdot^{***})$

and $\exists Z (Z \text{ is an infinite bad sequence in } \mathcal{T}(\cdot^{***}) \wedge \forall i < lh(\sigma) ((\sigma)_i = (Z)_i))$,

and $\psi(\sigma)$ as

σ is a finite sequence of elements in $\mathcal{T}(\cdot^{***})$

and $\forall Y [(Y \text{ is an infinite bad sequence in } \mathcal{T}(\cdot^{***}))$

$\rightarrow \forall i < lh(\sigma) [\forall k < i ((\sigma)_k = (Y)_k) \rightarrow (\sigma)_i \leq (Y)_i]]$,

Note that $(\sigma)_i \leq (Y)_i$ is interpreted as the inequality relation between natural numbers and not between elements of $\mathcal{T}(\cdot^{***})$. Using $(\Pi_1^1 - CA_0)^-$, there exists a set S such that $\sigma \in S \leftrightarrow \chi(\sigma) \wedge \psi(\sigma)$. Choose now two arbitrary elements s, s' in S . We want to prove that either s is an initial segment of s' or s' an initial segment of s . Assume there is an index $i < \min\{lh(s), lh(s')\}$ such that $(s)_i < (s')_i$ and $\forall k < i, (s)_k = (s')_k$. The case $(s)_i > (s')_i$ can be handled in a similar way. Note that $(s)_i < (s')_i$ is seen as an inequality between natural numbers and not between elements of $\mathcal{T}(\cdot^{***})$. Because s is in S , s can be extended to an infinite bad sequence Y in $\mathcal{T}(\cdot^{***})$. This, however, contradicts $\psi(s')$ because $(Y)_k = (s)_k = (s')_k$ for all $k < i$, but $(Y)_i = (s)_i < (s')_i$.

If $s \in S$, one can show by minimization in RCA_0 , that there is a $z \in \mathcal{T}(\cdot^{***})$ such that $s^\frown(z) \in S$. Therefore, there exists an infinite sequence $(s_i)_{i < \omega}$ in $\mathcal{T}(\cdot^{***})$ such that $S = \{(s_0, \dots, s_i) : i < \omega\}$.

Now, define $subS$ as the set of all pairs (i, t) such that t is a leaf of s_i . Remark that $subS$ is definable in RCA_0 , because

$$\begin{aligned} (i, t) \in subS &\Leftrightarrow \exists \sigma (\sigma \in S \text{ and } lh(\sigma) = i + 1 \text{ and } t \text{ is a leaf of } \sigma_i) \\ &\Leftrightarrow \forall \sigma ((\sigma \in S \text{ and } lh(\sigma) = i + 1) \rightarrow t \text{ is a leaf of } \sigma_i). \end{aligned}$$

On $subS$, we define the following ordering: $(i, t) \leq (j, t') \Leftrightarrow t \leq_{\mathcal{T}(\cdot^{***})} t'$. With this ordering $subS$ is a quasi-order. We want to prove that it is a well-quasi-order.

Assume that this is not true. This implies the existence of an infinite bad sequence $((n_i, s'_i))_{i < \omega}$ in $subS$. This implies $s'_i \leq s_{n_i}$ for all i . Construct now an infinite subsequence H such that $(n_i)_{i < \omega}$ is strictly increasing and the first element of H is (n_0, s'_0) . So $H = \{(i, (n_{j_i}, s'_{j_i})) : i < \omega\}$ with $j_0 = 0$. This is possible in RCA_0 , because the number of leaves of an element in $\mathcal{T}(\cdot^{***})$ is finite.

Construct now recursively a set S' such that all elements that lie in S' have the form $(\sigma_0, \dots, \sigma_i)$ such that if $i < n_0$, the element σ_i is equal to s_i and if $i \geq n_0$, then σ_i is equal to $s'_{j_i - n_0}$. The existence of S' is however in contradiction with the definition of S . This is because $\psi((s_0, \dots, s_{n_0}))$, $s'_{j_0} = s'_0 < s_{n_0}$ as a natural number and $(\sigma_i)_{i < \omega}$ is an infinite bad sequence. Take for example $k < l < \omega$. If $k < l < n_0$, then $\sigma_k \not\leq \sigma_l$ follows from $s_k \not\leq s_l$. If $n_0 \leq k < l$, then $\sigma_k \not\leq \sigma_l$ follows from $s'_{j_k - n_0} \not\leq s'_{j_l - n_0}$. Assume $k < n_0 \leq l$. Then $\sigma_k \not\leq \sigma_l$ follows from the fact that otherwise $s_k \leq s'_{j_l - n_0} \leq s_{n_{j_l - n_0}}$ with $k < n_0 = n_{j_0} \leq n_{j_l - n_0}$, a contradiction. So we conclude that $\text{sub}S$ is indeed a well-quasi-order.

By Lemma 6.3, we obtain that $(\text{sub}S)^{***}$ is a well-quasi-ordering. Now, look at the infinite sequence $(s_i)_{i < \omega}$ in $\mathcal{T}(\cdot^{***})$. Rewrite every element $\times s_i$ to an element in $(\text{sub}S)^{***}$ and call it \overline{s}_i . For example, if $s_i = \circ((s_1^1, \dots, s_{n_1}^1), \dots, (s_1^k, \dots, s_{n_k}^k))$, then \overline{s}_i is equal to $((i, s_1^1), \dots, (i, s_{n_1}^1), \dots, ((i, s_1^k), \dots, (i, s_{n_k}^k)))$. Because we know that $(\text{sub}S)^{***}$ is a well-quasi-order, there exist two indices $i < j$ such that $\overline{s}_i \leq_{(\text{sub}S)^{***}} \overline{s}_j$. Therefore, $s_i \leq_{\mathcal{T}(\cdot^{***})} s_j$, a contradiction. \square

Remark that the previous proof can actually be simplified. We wrote it however in this way, so that the reader can see what is needed to prove the well-partial-orderedness of $\mathcal{T}(W)$: the proof can be carried out in $RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^- + \forall X(X \text{ is a wpo} \rightarrow X^* \text{ is a wpo})$. This theory is the same as $ACA_0 + (\Pi_1^1 - CA_0)^-$. A general guideline is that for a natural theory $T \supseteq RCA_0$ with $RCA_0 \vdash T \leftrightarrow \forall X(X \text{ is a wpo} \rightarrow W(X) \text{ is a wpo})$, the following holds

$$T + (\Pi_1^1(\Pi_3^0) - CA_0)^- \vdash \mathcal{T}(W) \text{ is a wpo}$$

and

$$T + (\Pi_1^1(\Pi_3^0) - CA_0)^- \not\vdash \mathcal{T}(W') \text{ is a wpo},$$

where W' is the ‘limit’ of W . The notion ‘limit’ is of course very vague. For example a binary tree in $\mathcal{B}(X)$ can be seen in some sense as an element of X^{***} and a finite sequence in X^* can be seen as an element of X^n . Therefore, $\mathcal{B}(\cdot)$ is the limit of \cdot^{***} and \cdot^* is the limit of \cdot^n . We do not give a specific definition of the notion ‘limit’ as this is only a guideline and not a real theorem.

6.2 Independence results for $RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-$

We know that $|RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-| = \vartheta(\Omega^\omega)$. Furthermore, one can easily show as in Theorem 4.3 that $RCA_0 \vdash \mathcal{T}(X^*) \text{ is a wpo} \rightarrow WF(\vartheta(\Omega^\omega))$. Therefore, $RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^- \not\vdash \mathcal{T}(X^*) \text{ is a wpo}$.

Again, as in the ACA_0 -case, one can search for provability results: $\mathcal{T}(X^*)$ can be seen as the limit of $\mathcal{T}(X^n)$. Therefore, it would be interesting if $RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^- \vdash \text{'}\mathcal{T}(X^n) \text{ is a wpo'}$. However, the theory RCA_0 does not prove $\forall X(\text{'}\mathcal{T}(X) \text{ is a wpo'}$ \rightarrow $\text{'}\mathcal{T}(X^n) \text{ is a wpo'}$) (see [4] for more information). The theorem $\forall X(\text{'}\mathcal{T}(X) \text{ is a wpo'}$ \rightarrow $\text{'}\mathcal{T}(X^n) \text{ is a wpo'}$) is however provable in $RCA_0 + CAC$, where CAC is the principle saying that every infinite sequence in a partial order has a subsequence that is either a chain or an antichain. This implies that $RCA_0 + CAC + (\Pi_1^1(\Pi_3^0) - CA_0)^- \vdash \text{'}\mathcal{T}(X^n) \text{ is a wpo'}$.

Acknowledgements

The first author wants to thank his funding organization Fellowship of the Research Foundation - Flanders (FWO). The second author's research was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. The authors also wish to thank Leszek Kołodziejczyk for his helpful comments and Ryota Akiyoshi for his fruitful discussions with the first author.

References

- [1] Becker, T., Weispfenning, V.: Gröbner bases, *Graduate Texts in Mathematics*, vol. 141. Springer-Verlag, New York (1993). A computational approach to commutative algebra, In cooperation with Heinz Kredel
- [2] Buchholz, W., Feferman, S., Pohlers, W., Sieg, W.: Iterated inductive definitions and subsystems of analysis: recent proof-theoretical studies, *Lecture Notes in Mathematics*, vol. 897. Springer-Verlag, Berlin (1981)
- [3] Buchholz, W., Schütte, K.: Proof theory of impredicative subsystems of analysis, *Studies in Proof Theory. Monographs*, vol. 2. Bibliopolis, Naples (1988)
- [4] Cholak, P., Marcone, A., Solomon, R.: Reverse mathematics and the equivalence of definitions for well and better quasi-orders. *J. Symbolic Logic* **69**(3), 683–712 (2004)
- [5] Clote, P.: The metamathematics of Fraïssé's order type conjecture. In: Recursion theory week (Oberwolfach, 1989), *Lecture Notes in Math.*, vol. 1432, pp. 41–56. Springer, Berlin (1990)

- [6] Girard, J.Y.: Proof theory and logical complexity, *Studies in Proof Theory. Monographs*, vol. 1. Bibliopolis, Naples (1987)
- [7] Higman, G.: Ordering by divisibility in abstract algebras. *Proc. London Math. Soc.* (3) **2**, 326–336 (1952)
- [8] Hirst, J.L.: Reverse mathematics and ordinal exponentiation. *Ann. Pure Appl. Logic* **66**(1), 1–18 (1994)
- [9] Hirst, J.L.: Reverse mathematics and rank functions for directed graphs. *Arch. Math. Logic* **39**(8), 569–579 (2000)
- [10] Jäger, G., Strahm, T.: Bar induction and ω model reflection. *Ann. Pure Appl. Logic* **97**(1-3), 221–230 (1999)
- [11] de Jongh, D.H.J., Parikh, R.: Well-partial orderings and hierarchies. *Nederl. Akad. Wetensch. Proc. Ser. A* **80** = *Indag. Math.* **39**(3), 195–207 (1977)
- [12] Kruskal, J.B.: The theory of well-quasi-ordering: A frequently discovered concept. *J. Combinatorial Theory Ser. A* **13**, 297–305 (1972)
- [13] Laver, R.: On Fraïssé’s order type conjecture. *Ann. of Math.* (2) **93**, 89–111 (1971)
- [14] Marcone, A.: On the logical strength of Nash-Williams’ theorem on transfinite sequences. In: *Logic: from foundations to applications* (Staffordshire, 1993), *Oxford Sci. Publ.*, pp. 327–351. Oxford Univ. Press, New York (1996)
- [15] Pohlers, W.: Proof theory, *Lecture Notes in Mathematics*, vol. 1407. Springer-Verlag, Berlin (1989). An introduction
- [16] Pohlers, W.: Proof theory. Universitext. Springer-Verlag, Berlin (2009). The first step into impredicativity
- [17] Rathjen, M., Weiermann, A.: Proof-theoretic investigations on Kruskal’s theorem. *Ann. Pure Appl. Logic* **60**(1), 49–88 (1993)
- [18] Schmidt, D.: Well-partial orderings and their maximal order types. *Habilitationsschrift*, Heidelberg (1979)
- [19] Schütte, K.: Proof theory. Springer-Verlag, Berlin-New York (1977). Translated from the revised German edition by J. N. Crossley, *Grundlehren der Mathematischen Wissenschaften*, Band 225

- [20] Schütte, K., Simpson, S.G.: Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen. *Arch. Math. Logik Grundlagen*. **25**(1-2), 75–89 (1985)
- [21] Simpson, S.G.: Nonprovability of certain combinatorial properties of finite trees. In: Harvey Friedman’s research on the foundations of mathematics, *Stud. Logic Found. Math.*, vol. 117, pp. 87–117. North-Holland, Amsterdam (1985)
- [22] Simpson, S.G.: Subsystems of second order arithmetic, second edn. *Perspectives in Logic*. Cambridge University Press, Cambridge (2009)
- [23] Van der Meeren, J., Rathjen, M., Weiermann, A.: Well-partial-orderings and the big Veblen number. *Archive for Mathematical Logic* pp. 1–38 (2014). doi: 10.1007/s00153-014-0408-5
- [24] Weiermann, A.: Proving termination for term rewriting systems. In: Computer science logic (Berne, 1991), *Lecture Notes in Comput. Sci.*, vol. 626, pp. 419–428. Springer, Berlin (1992)
- [25] Weiermann, A.: Bounds for the closure ordinals of essentially monotonic increasing functions. *J. Symbolic Logic* **58**(2), 664–671 (1993)
- [26] Weiermann, A.: An order-theoretic characterization of the Schütte-Veblen-hierarchy. *Math. Logic Quart.* **39**(3), 367–383 (1993)
- [27] Weiermann, A.: A computation of the maximal order type of the term ordering on finite multisets. In: Mathematical theory and computational practice, *Lecture Notes in Comput. Sci.*, vol. 5635, pp. 488–498. Springer, Berlin (2009)